

GenCore version 5.1.6
Cpymright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:23 ; Search time 4075.7 Seconds
(without alignments)
10068.203 Million cell updates/sec

Title: US-(9-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atgagcgaggataatgctta.....tgcacatcggaacatctg 1410

Scoring table: IDENTITY_NUC
Gapcp 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1410	100.0	2294	8	AF105034	AF105034 Arabidops
2	1397.4	99.1	3004	8	AF069468	AF069468 Arabidops
3	1397.4	99.1	103504	8	AFAC021640	AC021640 Arabidops
4	307.8	21.8	1175	8	AFNASC5D	X90454 A.thaliana
5	307	21.8	1163	8	AFNASC5D	AY088527 Arabidops
6	202.8	14.4	179428	8	AF003214	AP003214 Oryza sat
7	186.2	13.2	1155	8	AF081794	AF081794 Nicotiana
8	178.2	12.6	1148	8	AF099969	AF099969 Nicotiana
9	164.2	11.6	1381	6	AX155063	AX155063 Sequence
10	92.6	6.6	293	11	G71681	G71681 A85715834FM
11	65.8	4.7	310	11	G71612	G71612 A62988234FM
12	57.6	4.1	840	8	CNS0180K	AL110675 Botrytis
13	57.4	4.1	696	8	CNS018D5	AL114081 Botrytis
14	56.6	4.0	127709	2	CNS057C1	AL352983 Homo sapi
15	55.8	4.0	52359	2	AC010772	AC010772 Homo sapi
16	55.6	3.9	7218	6	I66494	I66494 Sequence 14
17	55.6	3.9	166050	2	AC013817	AC013817 Homo sapi
18	55.6	3.9	177008	2	AC011175	AC011175 Homo sapi
19	55.2	3.9	221133	2	AC122896	AC122896 Mus muscu
20	54.8	3.9	6485	6	AX251313	AX251313 Sequence
21	54.8	3.9	6485	6	AX346710	AX346710 Sequence
22	54.4	3.9	224448	2	PFMALP4	AL035477 Plasmodiu
23	54.2	3.8	6015	6	AX345578	AX345578 Sequence
24	54	3.8	7008	6	AX251285	AX251285 Sequence
25	54	3.8	114897	2	AP003624	AP003624 Oryza sat
26	53.8	3.8	210105	2	AC125948	AC125948 Rattus no
27	53.6	3.8	107739	2	AC116979	AC116979 Dictyoste
28	53.6	3.8	170568	2	AC107605	AC107605 Rattus no
29	53.4	3.8	83440	2	AC024285	AC024285 Homo sapi
30	53.4	3.8	137216	2	AC055869	AC055869 Homo sapi
31	53.4	3.8	273807	2	AC025421	AC025421 Homo sapi
32	53.2	3.8	131271	2	AC015927	AC015927 Homo sapi
33	53	3.8	152878	3	CEX18010A	AL034393 Caenorhab
34	52.8	3.7	64612	2	AC079333	AC079333 Homo sapi
35	52.8	3.7	74119	2	AC036177	AC036177 Homo sapi
36	52.8	3.7	147962	2	AC130632	AC130632 Rattus no
37	52.8	3.7	156550	2	AC015830	AC015830 Homo sapi
38	52.6	3.7	3364	14	TSP418778	AJ418778 Tomato sp
39	52.6	3.7	143585	2	AC013349	AC013349 Homo sapi
40	52.2	3.7	14867	3	AE001398	AE001398 Plasmodiu
41	52	3.7	9539	6	AX277889	AX277889 Sequence
42	52	3.7	9539	6	AX323566	AX323566 Sequence
43	52	3.7	163547	2	AC111523	AC111523 Rattus no
44	52	3.7	265985	2	AC087226	AC087226 Mus muscu
45	51.8	3.7	1141	6	AX083744	AX083744 Sequence

ALIGNMENTS

RESULT 1
AF105034
LOCUS AF105034
DEFINITION Arabidopsis thaliana delta7 sterol C-5 desaturase (STL) gene,
complete cds.
ACCESSION AF105034
VERSION AF105034.1 GI:5031218
KEYWORDS Arabidopsis thaliana.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2294)
AUTHORS Choe,S., Noguchi,T., Fujioka,S., Takatsuto,S., Tissier,C.P.,

Gregory,B.D., Ross,A.S., Tanaka,A., Yoshida,S., Tax,F.E. and Feldmann,K.A.	
The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol C-5 desaturation step leading to brassinosteroid biosynthesis Plant Cell 11 (2), 207-221 (1999)	
JOURNAL	99128169
MEDLINE	9927639
PUBMED	
REFERENCE	2 (bases 1 to 2294)
AUTHORS	Choe,S., Tanaka,A., Gregory,B.D. and Feldmann,K.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-NOV-1998) Plant Sciences, The University of Arizona, Forbes Hall 303, Tucson, AZ 85721, USA
FEATURES	Location/Qualifiers
source	1..2294
	/organism="Arabidopsis thaliana"
	/cultivar="Ws-2 (Wassilewskija-2)"
	/db_xref="taxon:3702"
	/chromosome="III"
	/map="2.2"
gene	43..1889
	/gene="STel"
	/note="DWARF7"
mRNA	join(43..368,736..1041,1396..1889)
	/gene="STel"
	/product="delta7 sterol C-5 desaturase"
CDS	join(143..368,736..1041,1396..1709)
	/gene="STel"
	/note="steroid/brassinosteroid biosynthetic enzyme"
	/codon_start=1
	/product="delta7 sterol C-5 desaturase"
	/protein_id="AAD38120.1"
	/db_xref="GI:5031219"
	/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWEPLPHFLTQWLR NYLAGTLLYFISGLKCFYIYILKINVLPKDAIPTIKAMRLQMFVAMKAMPYITLLP TVSEIYKQNTLSPKASIDFGLYFVIAIYLVFEFGIYMMHRELHDKLPYLYL HATHYIYKONTLSPAGLAFHPVDGLQAVPHVIALFIVPIHFTTHIGLLEWALWT ANTHDCIHGNIWPMVGAGYHTIIHTYIKHNYGHTITMDMWFGLRDLPLEEDNDKDS FKRAE"
variation	322
	/gene="STel"
	/note="dwf7-2 (stel-3) allele"
variation	1552
	/gene="STel"
	/note="dwf7-1 (stel-2) allele"
	/replace="a"
BASE COUNT	585 a 469 c 417 g 823 t
ORIGIN	
Query Match	100.0%; Score 1410; DB 8; Length 2294;
Best Local Similarity	100.0%; Pred. No. 4.7e-296;
Matches 1410; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGCGGCGGATATGCTTATCTGATGCAGTTGTGTGACGAAACCTCTTTTACAACCGA 60
Db	143 ATGCGGCGGATATGCTTATCTGATGCAGTTGTGTGACGAAACCTCTTTTACAACCGA 202
Qy	61 ATCGTTCTGAGTCATCTTTTCCGGCGGAATCTATGGGAACCCCTTACCTCAFTTCTCCAG 120
Db	203 ATCGTTCTGAGTCATCTTTTCCGGCGGAATCTATGGGAACCCCTTACCTCAFTTCTCCAG 262
Qy	121 ACATGGCTCGAAATACCTTCGCGGGAACCCCTACTATACCTTCATCTCCGGTTCCTCTGG 180
Db	263 ACATGGCTCGAAATACCTTCGCGGGAACCCCTACTATACCTTCATCTCCGGTTCCTCTGG 322
Qy	181 TGCCTTCTACATCTATTACCTTAAAAACAAGTTTACCTTCCCAAAGTCTCGACTTTCAC 240
Db	323 TGCCTTCTACATCTATTACCTTAAAAACAAGTTTACCTTCCCAAAGTCTCGACTTTCAC 382
Qy	241 TTTTGTATTCACATTGCTTAAATCGCTTTCTATGTTATCGATTTTTCATATTAAGGAAGA 300
Db	383 TTTTGTATTCACATTGCTTAAATCGCTTTCTATGTTATCGATTTTTCATATTAAGGAAGA 442

Db	2622	ACATCCATGACTGCATCCATCCATGCAACATCTG	2652	
RESULT 3				
ATAC021640				
LOCUS	ATAC021640	103904 bp	DNA	linear
DEFINITION	Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence, complete sequence.			PLN 24-JAN-2001
ACCESSION	AC021640			
VERSION	AC021640.7			GI:12408747
KEYWORDS	HTG			
SOURCE	Arabidopsis thaliana			
ORGANISM	Arabidopsis thaliana			
REFERENCE				
AUTHORS	Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B., Wu, D., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M.			
TITLE	Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence			
REFERENCE	Unpublished			
AUTHORS	Lin, X. and Kaul, S.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-JAN-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org			
COMMENT	3 (bases 1 to 103904)			
Lin, X.				
Direct Submission				
Submitted (24-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA				
On Jan 24, 2001 this sequence version replaced gi:12280770.				
Address all correspondence to: at@tigr.org				
BAC clone F16B3 is from Arabidopsis chromosome III and is near the molecular marker ml74.				
The orientation of the sequence is from SP6 to T7 end of the BAC clone.				
Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene (http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (http://www.tigr.org/tdb/at/at.html).				
Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted as tRNA-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by RepeatMasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).				
Location/Qualifiers				
1..103904				
/organism="Arabidopsis thaliana"				
/cultivar="Columbia"				
/db_xref="taxon:3702"				
/chromosome="III"				
/map="mi74"				
/clone="F16B3"				
<129..>979				
/gene="F16B3.1"				
/notes="predicted by genefinder"				
join(<129..241,586..>979)				
/gene="F16B3.1"				
join(129..241,586..979)				
/gene="F16B3.1"				
gene				
mRNA				
CDS				
FEATURES				
source				
1..103904				
/organism="Arabidopsis thaliana"				
/cultivar="Columbia"				
/db_xref="taxon:3702"				
/chromosome="III"				
/map="mi74"				
/clone="F16B3"				
<129..>979				
/gene="F16B3.1"				
/notes="predicted by genefinder"				
join(<129..241,586..>979)				
/gene="F16B3.1"				
join(129..241,586..979)				
/gene="F16B3.1"				
gene				
mRNA				
CDS				
note="hypothetical protein"				
/codon_start=1				
/protein_id="AAF32445.1"				
/db_xref="GI:6957701"				
/translation="MEKDEGSPLENEAGLLCYDDEYQGGPAPKQFRLYLSEIG ELQIGDEDDVISLADLYGAETAGYGCWENYEVRLKGLGLYTLGRHGVPTTKY AVNTPTSDDESALCAAEFFQDRDSVTKLLSDMHICDARPFVDYLPNSQFKSSGPGE SVFTCFSG"				
complement(<3079..>4265)				
/gene="F16B3.2"				
/note="identical to COL2 GB:AAB67879 from [Arabidopsis thaliana]"				
complement(join(<3079..3438,3582..>4265))				
/gene="F16B3.2"				
complement(join(3079..3438,3582..4265))				
/gene="F16B3.2"				
/codon_start=1				
/product="COL2"				
/protein_id="AAF32446.1"				
/db_xref="GI:6957702"				
/translation="MLKESNESGTWARACTCRSACTVYCEADSAIYCTCTDARVH AANRVASRHRVRCOCESAPAEFLCADAASLCTACDAEIHSAANLARRHORVPL PLSANSCSSMAPSETDADNEDDDREVASLLPNPKGNIGNONNGFLGVEYLDLDVYS SSMNDFDNQITHORSGDGVPLQVEESTSHLQSOONFOLGNGFSSGSAHYN NSLKDNLNHSASVSSMDISVVPSTASDITVOHPRTFKETIDLSGPPQTVVQOOLTPM EREARYLRYREKKTKRKFDKTIYASRKAYAEIRPRIKGRFAKRIETEAEEEFSTS LMSCTGYGVPSF"				
5228..5255				
/rpt_family="(CAA)n"				
<5471..>7228				
/gene="F16B3.3"				
/note="predicted by genefinder"				
<5471..>7228				
/gene="F16B3.3"				
5471..7228				
/gene="F16B3.3"				
/note="hypothetical protein"				
/codon_start=1				
/protein_id="AAF32447.1"				
/db_xref="GI:6957703"				
/translation="MVPPLKLKIDFTQPRAGDSLGFKPGSTIRIGTRVRCNELAIKDA GISTKHLRIVSDSENIHDLGSSNGTILNSDTPVNLSGDEIKLGVEYTLIV NGSDVQAPQEHKLPFRPRNRNKLAAADPPDPPIESVQEKPRARGSKQENELP KTSRARKNLDDIADKEELDVEIEKVKARPRKNAGSAIAKEEVEVEKKRVG RPKNASSAITEEEVEEKGSRARRKNSIYQSKLEVEDTPKAVEISEKSR KRYTRSKQIENECEFLVEGDEKTRTSKTTEIGGSELEWLNQARKSRARR KMDPESKETRNDDAGCEVLKNCHEEDKNEAQEGCSGSDDICQEDDEKCDGSKR VQGVETELKSTVEGLKCTKEDGETENSODTERENENKEAQEGCSERSDKVE RVGGAKRVEQVEIELKSKSTVGEDDLNCTVRDGETENLQEECHDESCKVEE AGFATLDEKVGQGWNNKVERVEVYLEKMKLREWFDAIEVQLPKOTIETEKMIEPM RSKSMRVHRIAEQKEKGD"				
complement(<7759..>10518)				
/gene="F16B3.4"				
/note="predicted by genscan, contains Coesterase domain"				
complement(join(<7759..7920,8001..8091,8192..8335, 8796..8861,8964..9101,9182..9283,9433..9507,9638..9718, 10256..>10518))				
/gene="F16B3.4"				
complement(join(7759..7920,8001..8091,8192..8335, 8796..8861,8964..9101,9182..9283,9433..9507,9638..9718, 10256..10518))				
/gene="F16B3.4"				
/note="hypothetical protein"				
/codon_start=1				
/protein_id="AAF32448.1"				
/db_xref="GI:6957704"				
/translation="MQLSPERCPRMSENREANSSEMLLHGSNRLSPSEHVRVRY SGNSEDGSPRICROQSPGRCHPRAAETYLITRLSNLGLYLGLDLYIPTSDGLKP VYVFTGGAIIIGYKAWGLLGLAERDIIIVCLDYRNFPQGTISDMVSDAAQGSF VCNNISAFGDPNRIYLMGOSAGAHISSCALFQAIKESRGESISVSQIKAYFGLS GYVNLNLFHFNRLGRLYRIFLSIMEGESPKQFSPVRLKDLNVRKAAALPHIIL FGSADYSIPPEASKYTDALQAAVEKALVYKGTHTDILFLQDPLRGCKDELFDHI VSMHADDSDALRNDVAPRRKLVPFELLKLACRVSPF"				

gene	complement(11800..14456)	Db	62809	ACATGGCTCCGAAATACCTCCGGAACCCCTACTATATCTTCATCTCCGGTTTCCTCTGG	62868
mRNA	/gene="F16B3.5"				
	/complement(join(11800..11972,12355..12394,12527..12640,12750..12834,12919..13039,13133..13258,13559..13659,13745..13874,14237..14456))	Qy	181	TGCTTCTACATCTATTACCTTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTCAC	240
	/gene="F16B3.5"	Db	62869	TGCTTCTACATCTATTACCTTTAAATCAACGTTTACCTTCCCAAGGCTCTCGACTTTCAC	62928
CDS	/complement(join(11820..11972,12355..12394,12527..12640,12750..12834,12919..13039,13133..13258,13559..13659,13745..13874,14237..14413))	Qy	241	TTTTGTATTCATCTATTGCTTAATCGCTTCTATGTTATCGATTTTTCATTTT-AAAGGAAG	299
	/gene="F16B3.5"	Db	62929	TTTTGTATTCATCTATTGCTTAATCGCTTCTATGTTATCGATTTTTCATTTTAAAGGAAG	62988
	/note="unknown protein"				
	/codon_start=1	Qy	300	AGGTTTCTTCGTGTACTGTACAGTAATTTGGATTTGATGGATAGTTCATGTTGCAT	359
	/protein_id="AAF32449.1"				
	/db_xref="GI:6957705"	Db	62989	AGGTTTCTTCGTGTACTGTACAGTAATTTGGATTTGATGGATAGTTCATGTTGCAT	63048
	/translation="MAEGEDSQRKKIAAFAFYENDARWADYWNILIPPHMASRP	Qy	360	TTATTGATTATTGTGTCATATCTCCATCTAAGGATTTGAACAGTTAGTGGCTTATATAA	419
	EVDFHFRKTYQRIIDPLVVEPSTSSSSOSARPTASSTASSPANNPOVRSRNS	Db	63049	TTATTGATTATTGTGTCATATCTCCATCTAAGGATTTGAACAGTTAGTGGCTTATATAA	63108
	3SVPTSPGATTCATPSSMRWDEQTIQFSVNAWVFVIAVLAVLPLPKNLSNRYRL				
	SFMGTACSSLSLSLYGRPRANWQGLQYFQSIIVAAKDFIYIYCLTFTVSHCLK				
	FALLIPICRALEQVAKFELRNFRGKSTIYRKLEDPVWVESNTTLLNLSSQAEIAG				
	FLLIISLWORNLIQTFMYWQLLKLMYQAPVTAGYHQSTWSRIGRTVTPIIQYAPF	Qy	420	GTTTTGTGCAACCAATGAGAAGTCGTACATCTTTTGAAGTTGAATTTTCTACTTCCCAT	479
	LNTPEVAVQRWFR	Db	63109	GTTTTGTGCAACCAATGAGAAGTCGTACATCTTTTGAAGTTGAATTTTCTACTTCCCAT	63168
	<15001..>15660				
	/gene="F16B3.6"	Qy	480	TAAGTCCACATTAAATTTGTTTGTGAAGTGAATTTGCTTCCAGACATCTTTTCTG	539
	/note="predicted by genefinder"	Db	63169	TAAGTCCACATTAAATTTGTTTGTGAAGTGAATTTGCTTCCAGACATCTTTTCTG	63228
	<15001..>15660				
	/gene="F16B3.6"	Qy	540	CTTCTCTCAGACTCTGCTTGTAGTTTGAATCTTTTGTGCTGCTTGTGCTTCCAGATG	599
	/note="hypothetical protein"	Db	63229	CTTCTCTCAGACTCTGCTTGTAGTTTGAATCTTTTGTGCTGCTTGTGCTTCCAGATG	63288
	/codon_start=1				
	/protein_id="AAF32450.1"	Qy	600	TTCTTACAAATAAAGCTATGCTTTTGCATAATGTTTGTGCAATGAAGCTATGCCATGG	659
	/db_xref="GI:6957706"	Db	63289	TTCTTACAAATAAAGCTATGCTTTTGCATAATGTTTGTGCAATGAAGCTATGCCATGG	63348
	/translation="MSALRLRNANTPAPELDELSDQTPSESRLKROMSQRAMSNT				
	LTSANLSNLLPTCLLAFOLLTPVFTSNGVCDHATRELFATVLLFLAASCFVSFTD				
	SVKADDTGIVFGFTFGKMWVDYDPDGLGLPDLAKYRMRFVDMHATLSLVFEGAV	Qy	660	ACACTCTTCTTCCAACTGCTCCGAGAGATGATGAAGCTGGTGGACCAATGTTTGG	719
	ALRDKYITDCFPSPAEATKRVLDIVPVGVGWCSLLFENVPARRHGIGLYLTVGSVR	Db	63349	ACACTCTTCTTCCAACTGCTCCGAGAGATGATGAAGCTGGTGGACCAATGTTTGG	63408
	R				
	complement(<16445..>17870)				
	/gene="F16B3.7"	Qy	720	CTAGCATAGAGCAATTCGGCTGGATCTGTGATTTTGTATTTACATGCCCATCTATCTTGT	779
	/note="predicted by genscan"	Db	63409	CTAGCATAGAGCAATTCGGCTGGATCTGTGATTTTGTATTTACATGCCCATCTATCTTGT	63468
	complement(join(16445..16464,16657..17107,17220..>17870))				
	/gene="F16B3.7"	Qy	780	TCGTTGAGTTGGTATTTATTGGATGACAGAGCTTCATGACATTAAGCCCTCTCTATA	839
	/note="hypothetical protein"	Db	63469	TCGTTGAGTTGGTATTTATTGGATGACAGAGCTTCATGACATTAAGCCCTCTCTATA	63528
	/codon_start=1				
	/protein_id="AAF32451.1"	Qy	840	AGTATCTCCATGCCACCCATCATATCTACAAACAGCAGATACACTCTCTCCATTTGCCG	899
	/db_xref="GI:6957707"	Db	63529	AGTATCTCCATGCCACCCATCATATCTACAAACAGCAGATACACTCTCTCCATTTGCCG	63588
	/translation="MELPATKLRIGLVIFPLLLTIAPILYLFPGYPLIYKSTYKHL				
	SNSSSAISPSRYNHNSSSDSYKTEDSEPSYDNDYDTHDPKSSLSHNDRLS	Qy	900	GTAAGTGTTCAGTTGCTTCTTCTTTAGTTCTTCTTAAAGATTTGGTAGCATTTAGTTTC	959
	ISSNGHQVTPKRRKRRKRCDFSGEWIPNKPAPYNTTCRAIHEHQCIRY	Db	63589	GTAAGTGTTCAGTTGCTTCTTCTTCTTAAAGATTTGGTAGCATTTAGTTTC	63648
	GRPLDGLMKWRKPKEDLPDFEYELVIRGTRMAFVGDSVSRNHNVSILCLLSRV				
	EHPEGDSQEFNFQWKYKTYNETIATFTWTHLVRAETETGTPGPSNFYNLIDEPD				
	PTWASQIIGEDYIILISSQHFREFLEFDKRIKIGLCYCIPIPGYRNVGAFYRRLR	Qy	960	TTACAGAAAAGACTTTTCTCAGCAGCTGCTTGTACTCCAAATACATTTTGCATTCCTTA	1019
	TTFTKILGLENFKEVFLRFAPSFHFEHGEKMGAMRQS	Db	63649	TTACAGAAAAGACTTTTCTCAGCAGCTGCTTGTACTCCAAATACATTTTGCATTCCTTA	63708
	complement(17673..17778)				
repeat_region	/rpt_family="CAT"	Qy	1020	TCCATAAAGTAACCAAGAAAGGCTAGAAATATATAATGTCAGCTCCATTTACTTTCACAT	1079
repeat_region	complement(18048..18074)	Db	63709	TCCATAAAGTAACCAAGAAAGGCTAGAAATATATAATGTCAGCTCCATTTACTTTCACAT	63768
	/rpt_family="(TA)n"				
	Query Match 99.1%; Score 1397.4; DB 8; Length 103904;	Qy	1080	GTCAGAGAGACTCTTGACTTAACCAAGAGTTTAGATCTTTTGTGTTCTCTCTCGGCTCGG	1139
	Best Local Similarity 99.9%; Pred. No. 1.3e-293;	Db	63769	GTCAGAGAGACTCTTGACTTAACCAAGAGTTTAGATCTTTTGTGTTCTCTCTCGGCTCGG	63828
	Matches 1409; Conservative 0; Mismatches 1; Indels 1; Gaps 1;				
	1 ATGCGGCGGATAATGCTATCTCATGCTGAGTTCTTGACGAACCTCTTTTACACCGA 60	Qy	1140	ACTGATTGGAATGACGAGAAAGTTCTTTTATCTACTTCCCTGGAGTGTATCTTGGTTAAT	1199
		Db	63769	ACTGATTGGAATGACGAGAAAGTTCTTTTATCTACTTCCCTGGAGTGTATCTTGGTTAAT	63883
	61 ATCGTTTGTAGTCACTTTTGGCGGCGAATCTATGGAACCTTACCTTCATTTTCTCCAG 120	Qy	1200	CCAAGGATGTGACATCTAATATTAATCTTCACTTCCCTACGTTTGTGTTTACAGGCTTG	1259
		Db	63829	CCAAGGATGTGACATCTAATATTAATCTTCACTTCCCTACGTTTGTGTTTACAGGCTTG	1259
	62749 ATCGTTTGTAGTCACTTTTGGCGGCGAATCTATGGAACCTTACCTTCATTTCTCCAG 62808				
		Qy	121	ACATGGCTCCGAAATACCTCGCGGAACCCCTACTATCTATCTCCGGTTTCTCTCTG	180

```
Db 63889 CCAAGAGTGTGACATCAATATATTACTGTAACTTCCTTACGTTTTTTTACAGGGCTTG 63948
QY 1260 CATTTCACCCAGTAGAGGATACCTACCGCTGTACCGATGTGATAGCGCTGTTTATAG 1319
Db 63949 CATTTCACCCAGTAGAGGATACCTACCGCTGTACCGATGTGATAGCGCTGTTTATAG 64008
QY 1320 TGCCAATTCATTTCACAACCTCATATAGGTCCTTTTGTTCATGGAAGCGATATGACGGCGA 1379
Db 64009 TGCCAATTCATTTCACAACCTCATATAGGTCCTTTTGTTCATGGAAGCGATATGACGGCGA 64068
QY 1380 ACATCCATGACTGCATCCATGCGCAACATCTG 1410
Db 64069 ACATCCATGACTGCATCCATGCGCAACATCTG 64099

RESULT 4
ATTRASC5D 1175 bp mRNA linear PLN 13-FEB-1997
LOCUS A.thaliana mRNA for sterol-C5-desaturase.
DEFINITION X90454
ACCESSION
VERSION X90454.1 GI:1061037
KEYWORDS sterol-C5-desaturase gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1175)
Cachotte,D., Husselsstein,T., Bard,M., Lacroute,F. and Benveniste,P.
Isolation and characterization of an Arabidopsis thaliana cdna
encoding a delta 7-sterol-C-5-desaturase by functional
complementation of a defective yeast mutant
Plant J. 9 (3), 391-398 (1996)
JOURNAL 97077348
MEDLINE 8919915
PUBMED
REFERENCE 2 (bases 1 to 1175)
Gachotte,D.
Direct Submission
TITLE Submitted (12-JUL-1995) D. Gachotte, Institut de Biologie
JOURNAL Moleculaire Plantes, Enzymologie Moleculaire et Cellulaire,
Institut de botanique 28, rue Goethe, F- 67083 Strasbourg, FRANCE
FEATURES
source
Location/Qualifiers
1. .1175
/organism="Arabidopsis thaliana"
/variety="Landsberg erecta"
/db_xref="taxon:3702"
118. .963
/codon_start=1
/product="sterol-C5-desaturase"
/protein_id="CAA62079.1"
/db_xref="GI:1061038"
/db_xref="SPTREMBL:Q39208"
/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWEPFLHQTWLR
NYLQSLTIIISGFLKCVIYVIRKINVLPKDAIPTIKAMRLQMFVAMKMPWYLLP
TVSEMIERGKCFKASIGEGWILFYVYIYLVFVEGIYWMHRELHDLPLKYL
HATHYINKQNTSPFAGLAFVPGDILQAVPHVIAFLFVIFHFTHIGLLHMEAIWA
ANIHCQENIWPVAGYVHTIHTYKHYNYGHTYIWMDFGSLRDPLEEDDNKDS
FKKAE"
BASE COUNT 300 a 265 c 227 g 383 t
ORIGIN

Query Match: 21.8%; Score 307.8; DB 8; Length 1175;
Best Local Similarity 94.9%; Pred. No. 1.2e-56;
Matches 318; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 578 GTCTGTTTTCAGATCAATTCCTACAATAAGGCTATCGTTTGCATATGTTTGG 637
Db 328 GTTATCTCCCAAGATGCAATTCCTACAATAAGGCTATCGTTTGCATATGTTTGG 387
QY 638 GCAATGAAGGCTATGCGATGGTACACTCTCTTCCAACTGTCTCCGAGAGTATGATTGA 697
Db 388 GCAATGAAGGCTATGCGATGGTACACTCTCTTCCAACTGTCTCCGAGAGTATGATTGA 447
```

```
QY 698 CGTGTTGGACCAAAATGTTTCTAGCATAGACAGAAATTCGGCTGGATCTGTATTGTT 757
Db 448 CGTGTTGGACCAAAATGTTTCTAGCATAGACAGAAATTCGGTTGGATCTGTATTGTT 507
QY 758 TACATCGCCATCATCTCTGTTTCTGTTGAGTTTGGTATTATTGGATGCACAGAGCTT 817
Db 508 TACATCGCCATCATCTCTGTTTCTGTTGAGTTTGGTATTATTGGATGCACAGAGCTT 567
QY 818 CATGACATTAAGCCTCTCTATAGTATCTCCATGCCACCCATCATCTACACAAGCAG 877
Db 568 CATGACATTAAGCCTCTCTATAGTATCTCCATGCCACCCATCATCTACACAAGCAG 627
QY 878 AATACACTCTCTCCATTTGCCGGTAAAGTGTTC 912
Db 628 AATACACTCTCTCCATTTGCCGGCTTGCAATTTCA 662

RESULT 5
AY088527 1163 bp mRNA linear PLN 26-JUN-2002
LOCUS Arabidopsis thaliana clone 7564 mRNA, complete sequence.
DEFINITION AY088527
ACCESSION AY088527.1 GI:21407301
VERSION FLI_CDNA.
KEYWORDS thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1163)
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
2 (bases 1 to 1163)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1163)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Direct Submission
TITLE Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
JOURNAL Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
source
Location/Qualifiers
1. .1163
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="7564"
146. .991
/codon_start=1
/product="sterol-C5-desaturase"
/protein_id="AAM6060.1"
/db_xref="GI:21594971"
/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWEPFLHQTWLR
```

NVLGTLTYFTSGFLMCFYIYKINYLKPKDAIPTTKAMKLOMEVAMKAMPWYTLPL
TVSEMIERKXKGFASIGFEGWILYFYTAIYLVFEVFEFGYIWHHRELHDIKPKLYKIL
HATHIYNKQNTLSPFAGLAFPHVDGILOAPVHIALFIVPHTHTTGLGLFMEAIWT
ANIHCDCINGNWPVVGAGYHTIHTTYKHNHYGHTIWMWMEGSLRPLLEEDNDKDS
FKKAE"

BASE COUNT 291 a 263 c 223 g 384 t 2 others

Query Match 21.8%; Score 307; DB 8; Length 1163;

Best Local Similarity 94.3%; Pred. No. 1.8e-56;

Matches 316; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 578 GTCGTGTTTGTCAGATGCAATTCCTACATAAAGCGTATCGGTTTCCAAATGTTGTG 637

Db 356 GTTTACCTTCCAAAGATGCAATTCCTACATAAAGCGTATCGGTTTCCAAATGTTGTG 415

QY 638 GCAATGAJGGCTATGCCATGTACACTCTCTTCCAACTGCTCCGAGCATGATTCGAA 697

Db 416 GCAATGAJGGCTATGCCATGTACACTCTCTTCCAACTGCTCCGAGCATGATTCGAA 475

QY 698 CGTGGTTTGACCAATGTTTGTAGCATAGACGAATTCGGCTGATTCCTGATTTTGT 757

Db 476 CGTGGTTTGACCAATGTTTGTAGCATAGACGAATTCGGCTGATTCCTGATTTTGT 535

QY 758 TACATCGICATCTATCTGTTTGGTTGAGTTGGTATTTATGATGCACAGAGCGTT 817

Db 536 TACATCGICATCTATCTGTTTGGTTGAGTTGGTATTTATGATGCACAGAGCGTT 595

QY 818 CATGACAATAAGCCTCTATAAGTATCTCCATCCACCCATCATATCATCAACAGCAG 877

Db 596 CATGACAATAAGCCTCTATAAGTATCTCCATCCACCCATCATATCATCAACAGCAG 655

QY 878 AATACACICTCTCCATTTGCCGGTAAGTGTGTTTCA 912

Db 656 AATACACICTCTCCATTTGCCGGTTCATTTCA 690

RESULT 6

AP003214

LOCUS

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

AP003214

/translation="MESPPRSCSDHAAAIISVVGTRRRRLHIRMALAEARPGSPQLDSL
QSMRSHLGCADGDTADATACAGELHGLAWPRTPPCAAITRRRCRPPHPCGAAR
SAAARAMATGNNVSWPVLPVPPGAAVNPPLAPASAPSVPRRPPRPPSVAT
IGARGGPPERGAAGVGGRRAPAEVGPAPALGAVAAWCGGG"
complement(join(12004).12746,13413..13449))
/gene="OSJNb0083M16.4"
/complement(join(12004).12746,13413..13449))
/gene="OSJNb0083M16.4"
/complement(join(12004).12746,13413..13449))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64606.1"
/db_xref="GI:1528584"
/translation="MTVLGRKSSGLSVAVAMEELAPELIGPRISFSDGVVATAGA
ATPAAEDTSLVWVSSWRRLPEPEFDFAAAAADVADRLFAAGKLLPVPLPPVH
PKSPCKQOATSGGKQKPTSYQRRPGSWTSPFTRSCSVNSATTAAPRSGSGFSCPS
FPLMRSGAAGAGGGLGVSGGHHRRPQHKKAGATAAAYYGGSRNGSGHG
VRVSPVNPSTGTSNVNMLSYLLDCGKNKTRKGFGLNC"
complement(join(16789).17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNb0083M16.5"
complement(join(16789).17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNb0083M16.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64607.1"
/db_xref="GI:1528585"
/translation="MGRLLLVSLDRRRRVLVPSHCQVDPVVGWNRQAQCOH
DYHAGFEFPKLLLSLSPRPVRLPIARAKNSILMEGKHAVDLSSEERRG
GAPGSGSDEVDGAVLGGAESTEPGSSVDDHGVAAEAKNDDELEPLEPE
GVTAAGGAGTADSGGNEHLEVPKPMATRFAPRATDNATPVYRGMLSRSRK
NAGPTFVSAGPASPDPALSTSSSELVKSSENASADPADHAGSPGRNQDL
EAAKQSEARFADPAAAAAASREHPSRSKKQPRPERFIPGEAAARAKACRS
GALDRFITSQLNPGSPSTEWERTAADVVGGOGEPTTSDOPSCSIATSDSGPP
EPLPDREIRIVSLAVLGSLAVSMALVLEYFGESESPPSPDNQVIEKQV"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNb0083M16.6"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNb0083M16.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64608.1"
/db_xref="GI:1528586"
/translation="WISAEHPPPPDPVADGAAVEAAASARHQAALVSFPPLK
TWGSHVLCANVRAGDAIAAARRSPEOLDGKRELHLREVDSVVVADEAVAA
VAAPAAAAAHPKLRRRRNPAPASSPPERRAARAEALDRARFVTLSE
ETIEDVYATGARPRRRPRPVOKLDGIRNFVKIHFVSCIFPFPHGCPSPSP
TGCTPIVPGTHLSEELRVDTLYITGGPRGRGPQVCSASRTMKGKIHNVNEHHSF
LSPWNYLEPEVELDALRCQLSCSDCAPRPSRSSRSIYYNQHLQAMFELHQD
RLGGTGICKD"
complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNb0083M16.7"
complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNb0083M16.7"
/note="contains ESTs AU075676(E10683),AU075677(E10683)"
/codon_start=1
/product="similar to Mus musculus SURF-5 gene"
/protein_id="BAB64609.1"
/db_xref="GI:1528587"
/translation="MSKAGAAAGPTAAAAAAYQKQKTLQAKADVSSLVNFAAF
INIRVNDPVRNTQAFQMDMRGSRLLIHSDSLKLKLVSELKRTAIFSGLSITENV
RRIELSQOAEQTERLMRIGEAAGSLKELEAHYSSVVRTPPDE"
complement(42717..43520)
/gene="OSJNb0083M16.8"
complement(42717..43520)
/gene="OSJNb0083M16.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64610.1"
/db_xref="GI:1528588"
/translation="MSWTSPLKLFQFDVSEETABHEQYSAEPESAPVCGNGGGG

GGSDSSSSSTTTTTATAGDGRREYCYCCREFANSQALGGHQAHHKRRQQLKRA
QLQAAAAAGVAGGAAALYPRANPWSAFAPTPHLLGGEAGPTSWVYFSRAAFA
QAGQCOQHVSHGCVFARGAGPAVAASPAVFSVTPATSAALAAAGAPVYTTDEH
HGRRIHASSAASLARYPGAMVVAEPVATVPEDTCGLDQLQLSLAPAGI"
complement(join(57300..57361,57501..57585,57688..57936))
/gene="OSJNb0083M16.9"
complement(join(57300..57361,57501..57585,57688..57936))
/gene="OSJNb0083M16.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64611.1"
/db_xref="GI:1528589"
/translation="MSLATPASAARLSPLTSSPEPCRRRLLSAAAAPLRRLRLR
RIAVASVDPDAARPAEYPPWLIAGLNGPSKYHGRHNVPFEMVDRIARDEGTMNT
IQKSLILGIAETIVVEAIDLKQITLTFV"
complement(59958..63558)
/gene="OSJNb0083M16.10"
complement(59958..63558)
/gene="OSJNb0083M16.10"
/note="contains EST C26390(C12241)"
Query Match 14.4%; Score 202.8; DB 8; Length 179428;
Best Local Similarity 55.3%; Pred. No. 3.3e-34;
Matches 466; Conservative 0; Mismatches 357; Indels 19; Gaps 3;
QY 585 TTGCTTCAGATGCAATTCCTACAATAAAGGCTATGCGTTTGGCAAATGTTTGGCAATGA 644
DB 131331 TTGATGCAGATCTATACCTACATGAAGCTATGAAGAAGCAATAATTGTTGCATCAA 131390
QY 645 AGCTATGCGATGGTACACTCTCTTCCAACTGCTCCGAGAGTAGTATGTTGAACGGTGT 704
DB 131391 AGCTATGCTCTCTATGTTGCGCTTCCAACTTATCTGAGTACATGTTGAGAATGGAT 131450
QY 705 GGACCAAAATGTTTGTCTAGCATAGACGAATTCGGCTGGATTCGTATTTGTTTACATCG 764
DB 131451 GGACAGAGTGTATGTTATATATCAGTGAAGTGGTTGGCAATGTCACCTGGTTTATCTCG 131510
QY 765 CCATCTATCTGTTTTCGTTGAGTATTTGTTATTTATGATGTCACAGAGCTTCATGACA 824
DB 131511 CTTTATATCTATCTTTGTTGAGTATTTACTGGATTCGACAGAGAGTTGCATGACA 131570
QY 825 TTAAGCCTCTATAGATATCTCCATGCGCCACCATCATATCATACACAGCAGAAATACAC 884
DB 131571 TAAAGCCATGTACAAAGTACCTGCACACATACATATTTTACAACAAGGAGAAATACC 131630
QY 885 TCTCTCCATTGCGCGTAAAGTGT-----TTTCAGTTGTTGTTCTTTTAGTTCT 932
DB 131631 TATCACCATTTCGAGGTAGTATTTTATGGAGCCGTTTATATTTGCAATTCACITTTTA 131690
QY 933 TGTAAAGAGTATGTTAGCATTTAGTTCTTTACCAAGAAAGACTTTGTCAGCAGCTGCTGT 992
DB 131691 TGTTCGCTGTGCATATATATATGAGAACAGATTCCTATATTCGGAATGCTTATTTT 131750
QY 993 ACTCCAAATACATTTTGGCATTCCTTATCCATAAAGTAAACCAGAGGTAGATATATAT 1052
DB 131751 ACCATGCTTCCAGCTTTTATAGTATAATCTGTGTATAGTATATATATTTTCCCTGCTAA 131810
QY 1053 AAATGTCAGTGCATTTACTTTCATATATGTCAGAGACACTTCTGAC---TTAACACAGAT 1108
DB 131811 TTAATGTCGTTAAACACTAGGATACCTTATTTAAGGCTCCTAAACCTGTTACCCGTAGG 131870
QY 1109 TTAGATCTTTGTTGTTCTCTCTTCGCTCGGACTGATTTGGAATACGAGAAAGTCTTTT 1168
DB 131871 CCATACCATAATAAAAAAATGCGCTGCCATTCCTCAATGTACATTTTTTAAA---CAATA 131927
QY 1169 ATCTACTTCCCTGGAGTGTATCTTGGTAAATCCAAGATGTGACATCTATATATCTTGT 1228
DB 131928 CAATTTCTTTAGTTAACTCCCTTTGTTGCTGAGAAATGTTGATTTCTCCACCAGCTTCT 131987
QY 1229 AACTTCCCTTACGTTTGTGTTTACAGGGCTTGCAATTTCCACCAGTAGACGGGATCTCAG 1288
DB 131988 GATTTATGCTTGCATGCTTTCCAGGACTAGCATTCATCCACTGATGGGATTTTGCAA 132047

QY	1289	GCYGTACCCGATGTGATAGCGCTGTTTATAGTGCACATTCATTTCACAACTCATATAGGT	1348
Db	132048	GCCATACCCGATGTGTTTGGCGCTCTACTCTTATCCCAACACACTTTCAGGACACACATTCGT	132107
QY	1349	CTTTTGTTCATGAAGCGATATGACGGCGGAACATCCATGACTGCATCCCATGCGCAACATC	1408
Db	132108	CTTTTGTTCATGAAGCGCGTGTGGCAACATCAATCCATGACTGCATTCACGGCAAGTT	132167
QY	1409	TG TG 1410	
Db	132168	TG TG 132169	
RESULT 7			
AF081794			
LOCUS		1155 bp	mrna
DEFINITION		Nicotiana tabacum sterol-C5(6)-desaturase mRNA, complete cds.	
ACCESSION		AF081794	
VERSION		AF081794...	GI:4140397
KEYWORDS			
SOURCE			
ORGANISM			
		Nicotiana tabacum.	
		Nicotiana tabacum	
		Eukaryota Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Asteridae: euasterids I; Solanales; Solanaceae; Nicotiana.	
REFERENCE			
AUTHORS		1 (bases 1 to 1155)	
TITLE		Husselstein,T., Schaller,H., Gachotte,D. and Benveniste,P.	
JOURNAL		Delta7-sterol-C5-desaturase: molecular characterization and	
MEDLINE		functional expression of wild-type and mutant alleles	
PUBMED		Plant Mol. Biol. 39 (5), 891-906 (1999)	
REFERENCE			
AUTHORS		10344195	
TITLE		2 (bases 1 to 1155)	
JOURNAL		Benveniste,P.	
		Direct Submission	
		Submitted (30-JUL-1998) Plant Molecular Biology Institute, Centre	
		National de la Recherche Scientifique, 28 rue Goethe, Strasbourg	
		67084, France	
FEATURES			
source		Location/Qualifiers	
		1..1155	
		/organism="Nicotiana tabacum"	
		/cultivar="xanthi SH6"	
		/db_xref="taxon:4097"	
		/tissue_type="calli derived from leaf protoplasts"	
		103..918	
		/codon_start=1	
		/product="sterol-C5(6)-desaturase"	
		/protein_id="AAD04034.1"	
		/db_xref="GI:4140398"	
		/translation="MEDYLFQVEETSFYNRLVLGTMPESWGWGFLPHMLOGWLNRYI	
		GGVLLFISGFLWCFTYIHUKRNYIIPKDAIPSNKAMLLQLISVAKAMPWTCAFPSLS	
		EMYENKWTCKFARISDVGLSWLSYIAAILVIVFEGIYMMHMLHDIPKLYKYLHAT	
		HTIYNKQNTLSFAGLAFHPLDGLIQAVPHVALLPLVPMHFSTHIALIFLEALTANI	
		HDCIHGKVFVPMGAGYHTIHRRYRHNHYGTIIMDMWFGTLRDPVEEDAKKM"	
BASE COUNT		310 a 239 c 245 g 361 t	
ORIGIN			
Query Match		13.2%; Score 186.2; DB 8; Length 1155;	
Best Local Similarity		72.2%; Pred. No. 3.2e-30;	
Matches		242; Conservative 0; Mismatches 93; Indels 0; Gaps 0;	
QY	578	GTCTGTTTCGCTTCAGATGCAATTCCTACATAAAGCTATGCGTTTCCCAATCTTTGTG	637
Db	304	GTATATTTCTCTAAAGATGCCATCCCTTCTAACAAAGCAATGCTCTTGGCAATATCATGA	363
QY	638	GCAATGIAGGCTATGCCATGTCACACCTCTCTCCCACTGCTCCGAGAGTATGATTGAA	697
Db	364	GCCATGIAAGCGATGCCATGCTACTGTCGCCCTCCATCATTCTGAGTACATGATTGAA	423
QY	698	CGTGGTGGACAAAATGTTTTGCTAGCATAGACGAATTCGGCTCGATTCTGATTGTT	757
Db	424	AATGGAAGGACAAAATGTTTTGCAAGAAATAGTGTGATGCTGCTTTCCTGATTGATC	483

QY 698 CGTGGTTGGACCAAAATGTTTTGCTAGCATAGACGAATTCGGCTGGGATCTGATATTTGTT 757
DB 448 AATGGTGGACAAAATTTCTCAAGAATAGCGAGTTGGATGGCTTTCTACCTATT 507
QY 758 TACATCGCATCTATCTCTTTTCTGTTGAGTTGGTATTATTATTTGGATGCACAGAGCTT 817
DB 508 TATATCGCAGTTATTTGTTAGTAGAGAAATTCGGCAATTTACTGGATGATCGCGAGTTG 567
QY 818 CATGACATTAAGCTCTCTATAGTATCTCCATGCCACCATCATATCTACACACAGAG 877
DB 568 CATGACATTAAGCTCTCTACAAATATCTCCATGCTACACATCATATTTACACACAGAA 627
QY 878 AATACACTCTCTCCATTCGCGGTAAAGTGTTCFA 912
DB 628 AATACACTCTCCCATTTGCTGTTGGGTTCCA 662

RESULT 9
AX155063
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX155063
Sequence 109 from Patent WO0138484.
AX155063.1 GI:14536617
Physcomitrella patens.
Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 1381)
Lerchl, J., Renz, A., Ehrhardt, T., Reindl, A., Cirpus, P., Bischoff, F.,
Frank, M., Freund, A., Duwenig, E., Schmidt, R.M. and Reski, R.
Moss genes from Physcomitrella patens encoding proteins involved in
the synthesis of polyunsaturated fatty acids and lipids
Patent: WO 0138484-A 109 31-MAY-2001;
BASF Plant Science GmbH (DE)
Location/Qualifiers
1. .1381
/organism="Physcomitrella patens"
/db_xref="taxon:3218"
40. .1041
/note="PP004064012R"
/codon_start=1
/protein_id="CAC42667.1"
/db_xref="GI:14536618"
/translation="MASGAVNMVCAIAIVLMVMSLSICMSADVEVVNASFSSVVG
GAKTSGVVPAGSPYLAIFVEETRWYNDLVGLPSPVSRDIPHTLOTWLRNV
AGMLYFVSGGLWCLYVYSKGEHEFPAGDIPAKPEIMLIQWTKAMPVYGLPLS
EYMERGWTCKAREDEVGNLYVGLVIAVLAVVEGIYVWREHLIDKPLKHLAT
HHIYNKNTLSFPAFLAHPIDGILQACPHVIALFLPMHFFTHVELLCEGVWTTNI
HDCIDGNVNGIMGAGFHTIHTTYRHNYGHYTFVMDWLEGTLLRDPYERRKATARKSS"
BASE COUNT 323 a 331 c 351 g 376 t

Query Match 11.6%; Score 164.2; DB 6; Length 1381;
Best Local Similarity 68.7%; Pred. No. 1.8e-25;
Matches 226; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 584 TTGCTTCAGATCAATTCCTACAATAAGGCTATCGTTGCAAAATGTTGGCAATG 643
DB 421 TTCCCTGAGTGACATACCCGGAAGGAGCCATAATGCTCCAAATCTGGGTAATATG 480
QY 644 AAGGCTATGCCATGATGACTCTCTTCCCACTGCTCCGAGAGTATGATTGAACGTGGT 703
DB 481 AAGGCTATGCCATGATACACAGAGCTTCCCACTCTCTCCGAATATATGATTGAGCGGGG 540
QY 704 TGGACCAAAATGTTTCTAGCATAGACGAATTCGGCTGGATCTGATTTGTTTACATC 763
DB 541 TGGACCAAAATGTTTCTCGCGCTATCGAGGATGTTGGGTGGCTCACGATGATAGCCCTAGTC 600
QY 764 GCATCATCTCTGTTTCTGTTGAGTTGGTATTATTGATGCACAGAGAGCTTCATGAC 823
DB 601 ATCGCCTACTTGGCACTGTTGGAGTTGGTATCTATTGGATGCACAGAGAGCTTCAGAT 660

QY 824 ATTAAGCCTCTCTATAGTATCTCCATGCCACCATCATATCTACACACAGAGATACA 883
DB 661 ATTAAGCCTTTGTACAAACATCTGCTATGCTACCCACCATCTTACAATAAGCAAAACACG 720
QY 884 CTCCTCTCCATTTGCCGGTAAAGTGTTCFA 912
DB 721 CTATACCCGTTTGCAGGTTTGGCGTTCCA 749

RESULT 10
G71681/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

G71681
A85715834FMO17 maize leaf DNA Zea mays STS 08-JUN-2001
site.
G71681
G71681.1 GI:14333366
STS.
Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 293)
Yang, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
3' UTR sequences of maize genes
Unpublished (2001)
Contact: Schnable, P.S.
Schnable laboratory
Iowa State University
G405 Agronomy Hall, Ames, IA 50011, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Primer A: CTAAATAGCACACCAAGCC
Primer B: GATGGTATCTGCAAGCG
PCR Profile:
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR cycles: 31
Thermal cycler: Perkin Elmer TC
Protocol:
Template: 10-20 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total vol: 20 uL
Buffer:
MgCl2: 2 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.
Location/Qualifiers
1. .293
/organism="Zea mays"
/strain="DE811"
/db_xref="taxon:4577"
/clone_lib="maize leaf DNA"
/note="PCR products amplified from genomic DNA"
<1. .>293
BASE COUNT 77 a 51 c 81 g 84 t

Query Match 6.6%; Score 92.6; DB 11; Length 293;
Best Local Similarity 71.3%; Pred. No. 8.6e-10;
Matches 122; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1240 GTTTTGTGTACAGGCTTGCATTTCCACCCAGTAGCGGATACTTCAGGCTGTACCGCA 1299
DB 226 GTCTCCATTTGCTGGACTTGCATTTCCACCCAGTAGGTTGTTCTCGACCAATACCA 167

QY 1300 TGTGATAACGCTGTTTATAGTCCCAATTCATTCACAACTCATATAGGTCTTTTGTTCAT 1359
Db 166 TGTGTTACGCT 107
QY 1360 GGAAGCGTATGGAGCGGCGGACATCCATGATGATGATGATGATGATGATGATGATG 1410
Db 106 AGAGGGTGTGGGACAAACATCCAGGACTGATTCATGCGCAAGATATG 56

RESULT 11
G71612/c
LOCUS G71612 310 bp DNA linear STS 08-JUN-2001
DEFINITION A6298234:Zm017 maize leaf DNA Zea mays STS genomic, sequence tagged site.
ACCESSION G71612
VERSION G71612.1 GI:14333297
KEYWORDS STS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 310)
AUTHORS Yang, X.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
TITLE 3' UTR sequences of maize genes
JOURNAL Unpublished (2001)
COMMENT

Contact: Schnable, P.S.
Schnable laboratory
Iowa State University
G405 Agronomy Hall, Ames, IA 50011, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Primer A: CTAATAGACACCAAGCC
Primer B: GATGATCTCGAAGCG
PCR Profile:
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR cycles: 31
Thermal cycler: Perkin Elmer TC

Protocol:
Template: 10-20 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total vol: 20 uL
Buffer:
MgCl2: 2 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.
Location/Qualifiers
1. 310
/organism="Zea mays"
/strain="DE811"
/db_xref="taxon:4577"
/clone_lib="maize leaf DNA"
/note="PCR products amplified from genomic DNA"
STS
BASE COUNT 80 c 81 c 87 g 62 t
ORIGIN

Query Match 4.7%; Score 65.8; DB 11; Length 310;
Best Local Similarity 70.4%; Pred. No. 0.00056;
Matches 88; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1286 CAGCGCTTACGCGCATGTAGAGCGCTCTTATAGTCCCAATTCATTCACAACTCATATA 1345
Db 309 CAAGCAATGCGCGCATGTAGAGCGCTCTTCTCTCTCCCAAGCGCACTTCAGGACGCATC 250

QY 1346 GGTCTTTTGTTCATGGAAGCGATATGAGCGGAACATCCATGACTGCATCCATCCGCAAC 1405
Db 249 GCTCTCGTGTCTTGGAGGCGGTGGAGCAAAACATCCAGCACTGCATTCACGCGAAG 190
QY 1406 ATCTG 1410
Db 189 TTATG 185

RESULT 12
CNS0180K
LOCUS Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.
DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.
ACCESSION AL110675
VERSION AL110675.1 GI:5824962
KEYWORDS cDNA library; nitrogen deprivation.
SOURCE Botryotinia fuckeliana.
ORGANISM Botryotinia fuckeliana
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.
REFERENCE 1 (bases 1 to 840)
AUTHORS Bitton, F., Lewis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France
REFERENCE 2 (bases 1 to 840)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage; CP 5706 91057 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web: www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.
FEATURES
Location/Qualifiers
1. 840
/organism="Botryotinia fuckeliana"
/strain="T4"
/db_xref="taxon:40559"
/note="Genoscope sequence ID : W33E091"
BASE COUNT 151 a 23 c 79 g 539 t 48 others
ORIGIN

Query Match 4.1%; Score 57.6; DB 8; Length 840;
Best Local Similarity 46.3%; Pred. No. 0.028;
Matches 186; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
QY 235 TTTCACCTTTCTATTCACATATGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAA 294
Db 249 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 308
QY 295 GGAAGAGGGTTTCTTCGTGCTACTGTACAGTAAATTTGGATTTGATGATGATGATGATG 354
Db 309 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 368
QY 355 TGCATTTATGATTTATTTGTGCATATCTCCATCTCAAGGATGAACAGTTAGTGGCTTA 414
Db 369 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 428
QY 415 TATAAGTTTCTGCAACCAATGAGAAGCTGTACATCTTGAAGTTGAATTTCTACTGTG 474
Db 429 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 488
QY 475 CCATTTAAGTCCACTTAAATTTGTTTGAAGTGTATGCTACTTCTAGACACATCTTT 534
Db 489 TATTATTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 548
QY 535 TTCTGCTTCTCTGAGACT 594


```

Db 10196 TTCCTCCGATCGTATTGCAATATACNCCCTCCCTGTTNNNNNNNNNNNNNNNNNN 10137
QY 195 TTACCTTAAATCAACGTTTACCTCCCAAGGCTCGACTTTCACCTTTTGTATACACTA 254
Db 10136 NNNNNNTNNNNNNNTTTTNTTNNNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 10077
QY 255 TTGCTTAATCGCTTTCTAAGTATGATGATTTTCAATTTAAGAGAGGGTTCTTCGTGT 314
Db 10076 TTTTTCCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 10017
QY 315 ACTGTAACGTAATTTGGATTTGATGATGATGATGATGATGATGATGATGATGATG 374
Db 10016 NTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9957
QY 375 GCATATTTCCATCTAAGGATTTGAACAGTTAGTGGCTTATATATAAGTTTGTGCAACCA 434
Db 9956 TNNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 9897
QY 435 ATGAGAACTCGTACATCTTTGAAGTGTGAATTTTCTACCTGCAATTTAAGTCCACTAAAT 494
Db 9896 TTTTNTN'TTTTTTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 9837
QY 495 TGTGTGTGAAGTGTGCTACTTTCAGACACATCTTCTCTCTCTCTCTGAGACCT 554
Db 9836 NTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9777
QY 555 GTCTTAGTTGAAATCTTTTGGTCTGTTTGTCTTCAGAT 595
Db 9776 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 9736

```

```

RESULT 15
AC010772/c
LOCUS
DEFINITION Homo sapiens chromosome 15 clone 318_H_11 map 15, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC010772.1 GI:5917986
VERSION
KEYWORDS HTG; HTGS-PHASE0.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,C., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Galagan,A., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,K.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehocky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,B., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
* NOTE: this record contains 57 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for

```

```

* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 914: contig of 914 bp in length
* gap of unknown length
* 915 1826: contig of 912 bp in length
* gap of unknown length
* 1827 2677: contig of 851 bp in length
* gap of unknown length
* 2678 3628: contig of 951 bp in length
* gap of unknown length
* 3629 4580: contig of 952 bp in length
* gap of unknown length
* 4581 5603: contig of 1023 bp in length
* gap of unknown length
* 5604 6470: contig of 867 bp in length
* gap of unknown length
* 6471 7393: contig of 923 bp in length
* gap of unknown length
* 7394 8268: contig of 875 bp in length
* gap of unknown length
* 8269 9149: contig of 881 bp in length
* gap of unknown length
* 9150 10088: contig of 939 bp in length
* gap of unknown length
* 10089 10983: contig of 895 bp in length
* gap of unknown length
* 10984 11893: contig of 910 bp in length
* gap of unknown length
* 11894 12759: contig of 866 bp in length
* gap of unknown length
* 12760 13692: contig of 933 bp in length
* gap of unknown length
* 13693 14592: contig of 900 bp in length
* gap of unknown length
* 14593 15485: contig of 893 bp in length
* gap of unknown length
* 15486 16433: contig of 948 bp in length
* gap of unknown length
* 16434 17282: contig of 849 bp in length
* gap of unknown length
* 17283 18170: contig of 888 bp in length
* gap of unknown length
* 18171 19042: contig of 872 bp in length
* gap of unknown length
* 19043 19980: contig of 938 bp in length
* gap of unknown length
* 19981 20951: contig of 971 bp in length
* gap of unknown length
* 20952 21821: contig of 870 bp in length
* gap of unknown length
* 21822 22778: contig of 957 bp in length
* gap of unknown length
* 22779 23762: contig of 984 bp in length
* gap of unknown length
* 23763 24711: contig of 949 bp in length
* gap of unknown length
* 24712 25619: contig of 908 bp in length
* gap of unknown length
* 25620 26561: contig of 942 bp in length
* gap of unknown length
* 26562 27491: contig of 930 bp in length
* gap of unknown length
* 27492 28397: contig of 906 bp in length
* gap of unknown length
* 28398 29331: contig of 934 bp in length
* gap of unknown length
* 29332 30263: contig of 932 bp in length
* gap of unknown length
* 30264 31248: contig of 985 bp in length

```

* 31249 gap of unknown length
* 32162: contig of 914 bp in length
* 33089: contig of 927 bp in length
* 34038: contig of 949 bp in length
* 34936: contig of 898 bp in length
* 35972: contig of 1036 bp in length
* 36866: contig of 894 bp in length
* 37777: contig of 911 bp in length
* 38682: contig of 905 bp in length
* 39582: contig of 900 bp in length
* 40453: contig of 871 bp in length
* 41375: contig of 922 bp in length
* 42295: contig of 920 bp in length
* 43228: contig of 933 bp in length
* 44150: contig of 922 bp in length
* 45070: contig of 920 bp in length
* 46004: contig of 934 bp in length
* 46874: contig of 870 bp in length
* 47719: contig of 845 bp in length
* 48652: contig of 933 bp in length
* 49564: contig of 912 bp in length
* 50476: contig of 912 bp in length
* 51381: contig of 905 bp in length
* 52359: contig of 978 bp in length.

FEATURES
source

1..52359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="318_H_11"
/clone_lib="RPCI-11 Human Male BAC"
BASE COUNT 14555 a 7384 c 13014 g 7899 t 9507 others
ORIGIN

Query Match 4.0%; Score 55.8; DB 2; Length 52359;
Best Local Similarity 34.5%; Pred. No. 0.034;
Matches 138; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
QY 235 TTTCACCTTTGTAATGCGTTAAATCGCTTTCTATGATCGATTTTCAATTAA 294
Db 48642 TTTTNTNNNTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 48583
QY 295 GGAGAGGGTTTCCTGCTACTGACAGTAATTCGATTTGATGGAGTTCATGTT 354
Db 48582 TTTTNNNTNTTT 48523
QY 355 TGCATTTATGATTATTTGTCATATCTCCATCTAAGGATTAACAGTTCAGTGGCTTA 414
Db 48522 TTT 48463

QY 415 TATAAGTTTTTGTGCAACCAATGAGAAGTCGTACATCTTTTGAAGTTGAATTTCTACTTGG 474
Db 48462 TTTTNTTTTTNN 48403
QY 475 CCATTTAAGTCCACATTAAATTTGTTGTTGGAAGTGATTGCTACTTTTCACACACATTCITT 534
Db 48402 TTTTTTTTNTNTTTTTTTTTTTTTTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 48343
QY 535 TTCTGCTTCTCTGAGACTCTGCTTAGTTTGAATCTTTTTTTGGTCTGTTTGGCTTCAGA 594
Db 48342 TTTTTTTTTTTNN 48283
QY 595 TGCAATTCCTACAATAAAGGCTATGCGTTTGCAAAATGTTT 634
Db 48282 NNN 48243

Search completed: June 6, 2003, 15:14:08
Job time : 4084.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:23 ; Search time 520.302 Seconds
(without alignments)
10068.203 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322

Perfect score: 180

Sequence: 1 atgucgagcgaataatgctta.....tcattcgcgtttctctctgg 180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054610 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: cb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pln.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em_sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	180	100.0	1163	8	AY088527	Arabidops
2	180	100.0	2294	8	AF105034	Arabidops
3	180	100.0	3004	8	AF069468	Arabidops
4	180	100.0	103904	8	ATAC021640	Arabidops
5	178.4	99.1	1175	8	AF081794	A.thaliana
6	82	45.6	1155	8	AF081794	Oryza sat
7	67.6	37.6	179428	8	AF003214	AP003214
8	61.4	34.1	1148	8	AF099969	Nicotiana
9	54.8	30.4	1381	6	AX155063	Sequence
10	35.8	19.9	162212	2	AC113675	AC113675
11	35.8	19.9	174556	2	AC107404	AC107404
12	34.8	19.3	116788	9	AC007247	AC007247
13	33.2	18.4	186537	2	AC116209	AC116209
14	33	18.3	190998	9	AC019181	AC019181
15	32.6	18.1	10393	1	U32814	U32814
16	32.6	18.1	108795	2	AC124894	AC124894
17	32.6	18.1	156491	9	AL591720	AL591720
18	32.6	18.1	159914	9	AC027553	AC027553
19	32.6	18.1	168974	2	AC021299	AC021299
20	32.6	18.1	185822	2	AC073554	AC073554
21	32.6	18.1	196832	2	AL772211	AL772211
22	32.6	18.1	225957	9	AC069175	AC069175
23	32.4	18.0	31975	8	SPCC553	SPCC553
24	32.4	18.0	42776	3	CEF1504	CEF1504
25	32.4	18.0	100590	8	OSJN00191	OSJN00191
26	32.2	17.9	254752	2	AC127336	AC127336
27	32	17.8	120416	9	AC093915	AC093915
28	32	17.8	155223	9	AP001528	AP001528
29	32	17.8	175376	2	AC108223	AC108223
30	32	17.8	179491	2	AC020673	AC020673
31	32	17.8	184439	2	AP003085	AP003085
32	31.8	17.7	80568	9	AC002420	AC002420
33	31.8	17.7	107708	9	AL139803	AL139803
34	31.8	17.7	158073	9	AC069298	AC069298
35	31.8	17.7	164781	9	AC092440	AC092440
36	31.8	17.7	171031	2	AC104020	AC104020
37	31.8	17.7	174649	9	AC103586	AC103586
38	31.8	17.7	176030	2	AC100735	AC100735
39	31.8	17.7	180126	2	AC022074	AC022074
40	31.8	17.7	210000	2	AC006295	AC006295
41	31.8	17.7	232065	2	AC069425	AC069425
42	31.6	17.6	62088	2	AC104418	AC104418
43	31.6	17.6	66017	2	AC124262	AC124262
44	31.6	17.6	82642	9	AC103745	AC103745
45	31.6	17.6	156741	2	AC131312	AC131312

ALIGNMENTS

RESULT 1
AY088527
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AY088527
Arabidopsis thaliana clone 7564 mRNA, complete sequence.
AY088527
Arabidopsis thaliana
FLI_CDNA.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1163)
Haas,B.J., Volkovskiy,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

1163 bp mRNA linear PLN 26-JUN-2002

TITLE Full-length messenger RNA sequences greatly improve genome annotation

JOURNAL Genome Biol. (2002) In press

REFERENCE 2 (bases 1 to 1163)

AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.

TITLE Full-length cDNA from Arabidopsis thaliana

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1163)

AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the ws or laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones, Ceres, Inc, carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. .1163

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/clone="7564"

146. .991

/codon_start=1

/product="sterol-C5-desaturase"

/protein_id="AAM66060.1"

/db_xref="GI:21594971"

/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWELPHELOTWLR
NYLACTLLFISGLWCFFIYILKINYLKPKDAIPKAMRLQMFVAMKAMPWYLLP
TVSEMIERKNTKSPAGISGEFGLWILFYIYLVFVEGIYMHRELHDKPLIYKYL
HATHHLYKQNTLSFPAFLHPVDGILQAVPHVIALFIVPIHFTTHIGLLFMEALWT
ANHDCIHGNIPVMGAGYHTIHTTKYHNYGHTIIMDMFEGSLRDLLEEDNDKDS
FKKAE"

CDS

291 a 263 c 223 g 384 t 2 others

BASE COUNT 291 a 263 c 223 g 384 t 2 others

ORIGIN

Query Match 100.0%; Score 180; DB 8; Length 1163;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGGATAATCTTATCTGATGACGAGTTGTTGACGAACCTCTTTTACACCGA 60
|||||
Db 146 ATGGCGCGGATAATCTTATCTGATGACGAGTTGTTGACGAACCTCTTTTACACCGA 205
|||||

QY 61 ATGCTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCCCTTACTCTATTTCTCCAG 120
|||||

Db 206 ATGCTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCCCTTACTCTATTTCTCCAG 265
|||||

QY 121 ACATGCTCGGAATACCTTCGCGGGAACCCCTACTACTATCTATCTCCGGTTTCTCTGG 180
|||||

Db 266 ACATGCTCGGAATACCTTCGCGGGAACCCCTACTACTATCTATCTCCGGTTTCTCTGG 325
|||||

RESULT 2

AF105034 2294 bp DNA PLN 10-JUN-1999

LOCUS Arabidopsis thaliana delta7 sterol C-5 desaturase (STEL) gene, complete cds.

DEFINITION AF105034

ACCESSION AF105034

VERSION AF105034.1 GI:5031218

KEYWORDS

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 2294)

AUTHORS Choe, S., Noguchi, T., Fujioka, S., Takatsuto, S., Tissier, C.P., Gregory, B.D., Ross, A.S., Tanaka, A., Yoshida, S., Tax, F.E. and Feldmann, K.A.

TITLE The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol C-5 desaturation step leading to brassinosteroid biosynthesis

JOURNAL Plant Cell 11 (2), 207-221 (1999)

MEDLINE 99128169

PUBMED 9927639

REFERENCE 2 (bases 1 to 2294)

AUTHORS Choe, S., Tanaka, A., Gregory, B.D. and Feldmann, K.A.

TITLE Direct Submission

JOURNAL Submitted (06-NOV-1998) Plant Sciences, The University of Arizona, Forbes Hall 303, Tucson, AZ 85721, USA

FEATURES

Location/Qualifiers

1. .2294

/organism="Arabidopsis thaliana"

/cultivar="Ws-2 (Wassilewskija-2)"

/db_xref="taxon:3702"

/chromosome="III"

/map="2.2"

43. .1889

/gene="STEL"

/note="DWARF7"

join(43. .368,736. .1041,1396. .1889)

/gene="STEL"

/product="delta7 sterol C-5 desaturase"

join(143. .368,736. .1041,1396. .1709)

/gene="STEL"

/note="steroid/brassinosteroid biosynthetic enzyme"

/codon_start=1

/product="delta7 sterol C-5 desaturase"

/protein_id="AAD38120.1"

/db_xref="GI:5031219"

/translation="MAADNAYLMQFVDETSFYNRIVLSHLLPANLWELPHELOTWLR
NYLACTLLFISGLWCFFIYILKINYLKPKDAIPKAMRLQMFVAMKAMPWYLLP
TVSEMIERKNTKSPAGISGEFGLWILFYIYLVFVEGIYMHRELHDKPLIYKYL
HATHHLYKQNTLSFPAFLHPVDGILQAVPHVIALFIVPIHFTTHIGLLFMEALWT
ANHDCIHGNIPVMGAGYHTIHTTKYHNYGHTIIMDMFEGSLRDLLEEDNDKDS
FKKAE"

322

/gene="STEL"

/note="dwf7-2 (stel-3) allele"

/replace="a"

1552

/gene="STEL"

/note="dwf7-1 (stel-2) allele"

/replace="a"

585 a 469 c 417 g 823 t

BASE COUNT 585 a 469 c 417 g 823 t

ORIGIN

Query Match 100.0%; Score 180; DB 8; Length 2294;
Best Local Similarity 100.0%; Pred. No. 1.7e-44;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGGATAATCTTATCTGATGACGAGTTGTTGACGAACCTCTTTTACACCGA 60
|||||

Db 143 ATGGCGCGGATAATCTTATCTGATGACGAGTTGTTGACGAACCTCTTTTACACCGA 202
|||||

QY 61 ATGCTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCCCTTACTCTATTTCTCCAG 120
|||||

Db 203 ATGCTTCTGAGTCATCTTTTGGCGGGAATCTATGGAACCCCTTACTCTATTTCTCCAG 262
|||||

QY 121 ACATGCTCGGAATACCTTCGCGGGAACCCCTACTACTATCTATCTCCGGTTTCTCTGG 180
|||||

Db 263 ACATGCTCGGAATACCTTCGCGGGAACCCCTACTACTATCTATCTCCGGTTTCTCTGG 322
|||||

```
RESULT 3
AF069468
LOCUS
DEFINITION
Arabidopsis thaliana sterol-C5-desaturase gene, complete cds.
ACCESSION
AF069468
VERSION
AF069468.1
KEYWORDS
GI:4234767
SOURCE
Arabidopsis thaliana.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; euroids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS
Husselstein,T., Schaller,H., Gachotte,D. and Benveniste,P.
TITLE
Delta7-sterol-C5-desaturase: molecular characterization and
functional expression of wild-type and mutant alleles
JOURNAL
Plant Mol. Biol. 39 (5), 891-906 (1999)
MEDLINE
99273992
PUBMED
10344195
REFERENCE
2 (bases 1 to 3004)
AUTHORS
Husselstein,T., Schaller,H., Gachotte,D. and Benveniste,P.
TITLE
Direct Submission
JOURNAL
Submitted (30-MAY-1998) IBMP, 28, rue Goethe, Strasbourg 67083,
France
FEATURES
source
location/Qualifiers
1..3004
/organism="Arabidopsis thaliana"
/cultivar="columbia"
/db_xref="taxon:3702"
/tissue_type="leaves"
join(1242..1467,1836..2141,2496..2809)
/codon_start=1
/product="sterol-C5-desaturase"
/db_xref="GI:4234768"
/protein_id="AAD12944.1"
/translation="MAADNAYLMQFVDETSEYNRIVLSHLLPANLWELPHFLQTWLR
NYLAGTLFTSGFLWCFIYLYLKPDAIPTKAMLOFMVAMKMPWTLPL
TVSMTMERGTGCFASIGFEGWILYFYIAYLVFEFGIYWHRLHDIKPLYKYL
HATHIYNKONTLSPFAGLAPHPVDGILQAVPHVIALFIVPHETTHGLEFMEAINWT
ANIHDCHGNLWPVWGAGYTHIHTHTYKHNHYGHTIWNDFGSLRDLLEEDNDKDS
FKAAE"
CDS
865 .. 577 c. 490 g 1072 t
BASE COUNT
ORIGIN
Query Match
Best Local Similarity 100.08; Score 180; DB 8; Length 3004;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCGGATAATGCTATCTGATGTCAGTTTGTGACGAAACCTCTTTTACAAACCGA 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1242 ATGGCGCGGATAATGCTATCTGATGTCAGTTTGTGACGAAACCTCTTTTACAAACCGA 1301
QY 61 ATCGTTTGAGTCATCTTTTCGCGGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1302 ATCGTTTGAGTCATCTTTTCGCGGCGGAATCTATGGGAACCTTACCTCATTTTCTCCAG 1361
QY 121 ACATGGTTCGGAATACCTCGCGGGAACCTTACTATCTATCTCATCTCGGTTTCTCTGG 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1362 ACATGGTTCGGAATACCTCGCGGGAACCTTACTATCTATCTCATCTCGGTTTCTCTGG 1421
RESULT 4
ATAC021640
LOCUS
DEFINITION
Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence,
complete sequence.
ACCESSION
AC021640
VERSION
AC021640.7
KEYWORDS
GI:12408747
SOURCE
Arabidopsis thaliana.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; euroids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS
Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
Wu,D., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R.,
Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
TITLE
Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 103904)
AUTHORS
Lin,X. and Kaul,S.
TITLE
Direct Submission
JOURNAL
Submitted (18-JAN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 103904)
AUTHORS
Lin,X.
TITLE
Direct Submission
JOURNAL
Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 24, 2001 this sequence version replaced gi:12280770.
Address all correspondence to:at@tigr.org
COMMENT
BAC clone F16B3 is from Arabidopsis chromosome III and is near the
molecular marker mi74.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL (available by anonymous ftp
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
(http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
```


Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1175)
 Gachotte, D., Hüsselstein, T., Bard, M., Lacroix, F., and Benveniste, P. Isolation and characterization of an Arabidopsis thaliana cDNA encoding a delta 7-sterol-C-5-desaturase by functional complementation of a defective yeast mutant
 Plant J. 9 (3), 391-398 (1996)
 97077348
 8919915
 2 (bases 1 to 1175)
 Gachotte, D.
 Direct Submission
 Submitted (12-JUL-1995) D. Gachotte, Institut de Biologie Moléculaire, Enzymologie Moléculaire et Cellulaire, Institut de botanique 28, rue Goethe, F- 67083 Strasbourg, FRANCE
 Location/Qualifiers
 1. 1175
 /organism="Arabidopsis thaliana"
 /variety="Landsberg erecta"
 /db_xref="taxon:3702"
 118, 963
 /codon_start=1
 /product="sterol-C5-desaturase"
 /protein_id="CA62079.1"
 /db_xref="GI:1061038"
 /db_xref="SPTREMBL:O39208"
 /translation="MAADNAYLMQFVDETSEFYRNLVLSHLLPANIWEPLPHLOTWLR NYLAGTLLIISFGLWCFYIYRKNILVLPKDAIPKAMKLOFQVAMKAMPWTLLP TVSEMERQKTCFASIGFEGWILYFYAIYLVFVEFGIYWMHRLHDKPLRYKL HATHYIKQNTLSPPFAGLPHPLDGLQAVPHVIALFVPIHFTTHGLLEWAIWA ANTHDCIHGNIPWPMGAGYHTIHHTYKHNYGHYTIWMDWFMGSLRDLPEEDNDKDS FKAE"
 BASE COUNT 300 a 265 c 227 g 383 t
 ORIGIN
 Query Match 99.1%; Score 178.4; DB 8; Length 1175;
 Best Local Similarity 99.4%; Pred. No. 5.1e-44;
 Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ATGGCGCGGATATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACCGGA 60
 |||||
 Db 118 ATGGCGCGGATATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACCGGA 177
 |||||
 Oy 61 ATCGTTTGAGTATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCAG 120
 |||||
 Db 178 ATCGTTTGAGTATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTCTCCAG 237
 |||||
 Oy 121 ACATGGTCCGAATACCTGCGCGGAACCTACTATCTATCTCCGGTTTCTCTGG 180
 |||||
 Db 238 ACATGGTCCGAATACCTGCGCGGAACCTACTATCTATCTCCGGTTTCTCTGG 297
 |||||
 RESULT 6
 AF081794
 LOCUS
 DEFINITION
 Nicotiana tabacum sterol-C5(6)-desaturase mRNA, complete cds.
 AF081794
 VERSION
 AF081794.1 GI:4140397
 KEYWORDS
 Nicotiana tabacum.
 ORGANISM
 Nicotiana tabacum
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana. 1 (bases 1 to 1155)
 Hüsselstein, T., Schaller, H., Gachotte, D., and Benveniste, P. Delta7-sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles
 Plant Mol. Biol. 39 (5), 891-906 (1999)
 99273992
 10344195
 2 (bases 1 to 1155)

Benveniste, P.
 Direct Submission
 Submitted (30-JUL-1998) Plant Molecular Biology Institute, Centre National de la Recherche Scientifique, 28 rue Goethe, Strasbourg 67084, France
 Location/Qualifiers
 1. 1155
 /organism="Nicotiana tabacum"
 /cultivar="xanthi S96"
 /db_xref="taxon:4097"
 /tissue_type="calli derived from leaf protoplasts"
 103, 918
 /codon_start=1
 /product="sterol-C5(6)-desaturase"
 /protein_id="AAD04034.1"
 /db_xref="GI:4140398"
 /translation="MEDYILKQFVEETSEFYRNLVLTGPMPSWNGPLPHMLQGLRNLY GGVLLYFISGLWCFYIYHKNRYIYIPKDAIPSKMMLQISVAMKAMPWYCALPSLS EYMIENGWTKCFARISDVGLSYVYAIYLVFVEFGIYWMHMLHDKPLRYLHAT HHYIKQNTLSPPFAGLPHPLDGLQAVPHVIALFVPMHFTSHIALFLKALATANI HDCIHGKVEPVMGAGYHTIHHTYRHNHYGHYTIWMDWFMGTLRDPVEDAKKM"
 BASE COUNT 310 a 239 c 245 g 361 t
 ORIGIN
 Query Match 45.6%; Score 82; DB 8; Length 1155;
 Best Local Similarity 66.3%; Pred. No. 1.9e-14;
 Matches 118; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 Oy 3 GCGCGGATATGCTTATCTGATGCGAGTTGTTGACGAAACCTCTTTTACACCGGAAT 62
 |||||
 Db 96 GTCGAGATGAGGAGTACTTAAACACATTTGTGGAGGAGACATCGTTTACAAATCGGTT 155
 |||||
 Oy 63 GGTCTGAGTATCTTTTGGCGGCAATCTATGGAACCTTACCTCATTTTCTCCACAC 122
 |||||
 Db 156 GGTGTTGGTACATTCATGCGGCAATCATGTTGGGACCCTTCTCATGCTTCACGG 215
 |||||
 Oy 123 ATGCTCCGAAATACCTGCGGAAACCTACTATCTATCTATCTCCGTTTCTCTCG 180
 |||||
 Db 216 ATGCTCCGTAACATGTTGTTGTTCTCTTACTTATCTCCGTTTCTCTCG 273
 |||||
 RESULT 7
 AF003214
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC clone:OSUNBA0083M16.
 AF003214
 ACCESSION
 AF003214.3 GI:15528580
 VERSION
 AF003214.3
 KEYWORDS
 Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:OSUNBA0083M16.
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 1
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 Published only in database (2001)
 2 (bases 1 to 179428)
 Sakaki, T., Matsumoto, T. and Yamamoto, K.
 Direct Submission
 Submitted (19-FEB-2001) Takuji Sakaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Sep 7, 2001 this sequence version replaced gi:13620983.
 Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr

(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OSJNBa0083M16 clone has an overlap with P504H10 (DDBJ: AP003256) clone at the position 1 to 42,554 and with P0443D08 (DDBJ: AP003250) at the position 178,805 to 179,428 of 3' end. The sequence of this clone starts at the position 100,962 of P504H10 and ends at the position 624 of P0443D08. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

Source

```
1..179428
/organism="oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OSJNBa0083M16"
/complement(join(870..1087,1280..1364))
/gene="OSJNBa0083M16.1"
/complement(join(870..1087,1280..1364))
/gene="OSJNBa0083M16.1"
/notes="contains EST AU092803(C53725)"
/codon_start=1
/product="similar to Zea mays Bowman-Birk proteinase
inhibitors WIPI precursor"
/protein_id="BAB64603.1"
/db_xref="GI:1528581"
/translation="MKSSTLLVLAQAALINGIFAAVAKENAVGESKAIDINPQGLK
CCSCNCFSGSLYCTDDIVKCDPVCCKAVKTPVKMFKCTDFLGMGGPCKH"
5689..9051
/gene="OSJNBa0083M16.2"
5689..9051
/gene="OSJNBa0083M16.2"
/codon_start=1
/product="similar to Lycopersicon pimpinellifolium Cf-2
gene"
/protein_id="BAB64604.1"
/db_xref="GI:1528582"
/translation="MYSDFEPALFSLTSLQRLDLSMNSLGTSTTKDAEFDRLTSLT
HLNLSNGLDGQIPMGINKLINLVSLLDSKRYVNDNSDIFSNESDDEIIFTGDSYNHL
QESRLMSLVENSLNELYLVDHVDSTNVDDMCKTLAGSVPLQVLSLDGCSLNTPIH
HSLLRLSLTIVNLSQEGIAVLPDPEMGFANLTIVLRSNNLEGGWPPDKFEOLKN
LRILDSLNMLGLHLPKVTSLRLEGTNFTAKRISSSNFNMLKELGLEGLIS
KDFLSFGILNSLCHLELNSLGLSDLSNLISWIGAKNLTCLLSSEFSTKPPS
INFRNLSLWLCGNLRTPIMSAIGDLVDLQSLDMSNCTYSNMPISGINTNLSKL
YNSPFGFLPMPAALNKLKSLKSWFSCNCFETGMPSTIGNLTQKLTLEIAACRFSGP
IYISGOLKELRALFEGCNMSGRIPNSFTVNNMSKLIYGLPANLYSGKITPARLTVLPA
LFLDLFGNHFSGPIQEDAVPSYLSMLQSLTSLNETGPEPSPFFELTSLIALEIDLNN
LAGSVDSLFRKLRDLNLSNHLNLSVIMDEGNSNLSYLSLSELKELGLACCNITKF
PSILRLSDMSYLDLSCNKSIGNIPKWIWKRWSVVHLSNHLMTSMEVASYLLPF
NRHFTDLSSNLSLVLNLAHNFSGPPSCIMEQTYFNILNLRGNHFEGLMPTNVT
ISGNTPISCTNSLVLNLAHNFSGPPSCIMEQTYFNILNLRGNHFEGLMPTNVT
RCAFT IDLNGNK IEGRLPAGNCTYLEVLDLGNKADTPPSLGLSLNLRVLVR
SNRLYSGTYETEDSGDHPNQLIIDLASNFTSLGHPQFEKISMKKYNNTGETI
SHRSISDGYDTVTIISCKGFSMTFERILLTILADLSDNALEGISIPESVGLKLSLH
VLNLSHNAFSPGPIQIGITALELSDLSNWSIGEIPQELTNFTLFLVNLNSNOLE
KNSLPFRFATPENSSEYEGNAGLGDPLPKCASWSPPSAEPHVESSESHVDITVMFLFV
GVGFGVGAVALMKTSLWINRWFHSAVSQRQT"
/complement(join(9685..9961,10319..10662))
/gene="OSJNBa0083M16.3"
/complement(join(9685..9961,10319..10662))
/gene="OSJNBa0083M16.3"
```

```
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64605.1"
/db_xref="GI:1528583"
/translation="MESPRSCSDHAAISVWGTRRRLLHIRMAELAAAREPGSPQLDSL
QLSMQHLGCADDRTADATACGIELHGLAWPRTPPCAATTTCARRCASHPCGAAR
SSAARAMAIGCRNVSVYVLRPGAAVNNPPLAPASAPSPVRRPVRPPRRPSVAT
ISARGGGEERGAAGVGGRRAPAAVEGVPAPALPAAVAAWCGGGG"
/complement(join(12004..12746,13413..13449))
/gene="OSJNBa0083M16.4"
/complement(join(12004..12746,13413..13449))
/gene="OSJNBa0083M16.4"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64606.1"
/db_xref="GI:1528584"
/translation="MTVLGRKOSGLSVAVAPMEELAPELIGPRISFSHDGVVATAGA
ATPAARSDTSLVVMSSWRRLPEPEDFANAAADVAPADRLFAGCKLLPVPLPPVH
PKPSPCKQQATSGGGVKPTSVQRPGSWTSPTTRSCSVNSATTAAPRSGSGSFCPS
FPLMRSRAGSAAAGGGLGVGVGGGHHRPQHKKAGATAAAYYGGSRNGSGHG
VRVSPVINVPISGTSVMNMLSYLLCDCGKTKNRGFGLNC"
/complement(join(16789..17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNBa0083M16.5"
/complement(join(16789..17270,17559..17837,17983..18106,
18193..18498,18750..18902))
/gene="OSJNBa0083M16.5"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64607.1"
/db_xref="GI:1528585"
/translation="MGRLLLVSLDRRRRYIVLPSHCQGVDPVVCGRNAQRFQGH
DYHAGFERPKPLLSGSRPRVPRPLHPIARAKNSILMEEGKHAIVDLSEBERRG
GGAPGSGSDEVDGDAVGLGRGAESTEGSSAVDDAGHGVAAEAKNDLLEPLLEPE
GYTAVSGGGDGTASDGGENELHEVKNPMATREAFRATDANATPVSTYVGMLSBRK
NAGPTRFVSAGAPPADPTPALSTSSLSKSPENASVADPADHGASFGNGEQDL
EAAKQSEARFAPDAAAAAASHEPSKSKQPRPERFIEEGEAAAARAKACRS
EALDRFTTSQLNPSGPTSEWERVADVYGGGQECCTTSDPSCSIAISDSGPP
EPLDRRRRIYSVLAVLCVSLAVMALEFYIFGSESPSPSPNQVIFKQY"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNBa0083M16.6"
join(21657..22226,22293..22395,23106..23221,24403..24457,
25399..25589)
/gene="OSJNBa0083M16.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB64608.1"
/db_xref="GI:1528586"
/translation="MVISAEHPHPPPPDPVADGAAVEAAASARHQAALVSFFPSLK
TWGSHRVLCRCAHVNAGDAIAAARRSPQLDGVKRLLLHLREVDSYVADAEEVAAA
VAPAPAAAASRPWKLRSRRRPAVAPASAPSPERRAARAEALDRARFVTLTSE
ETEDVYVATGARPRRRPRRPVKQLDGPTRNFVKTHVPSCYFPVHGCRSPSPSSP
TCGPTIVPRGTHLSEELRWDTYLTGGPGRGRCQVCASRTKKGKTHNNVHHSF
LSPWNYLEPGEVELDALRCQLSCSDCAPRPRSSSRSIYINQHLQAMFFLHGQD
RLGGTGICND"
/complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNBa0083M16.7"
/complement(join(30012..30245,31517..31582,31671..31817))
/gene="OSJNBa0083M16.7"
/notes="contains ESTs AU075676(E10683).AU075677(E10683)"
/codon_start=1
/product="similar to Mus musculus SURF-5 gene"
/protein_id="BAB64609.1"
/db_xref="GI:1528587"
/translation="MSKAGAAGTAATAAAAAVOKTKLLLOKADADVSSLYDNAAF
INIARVNPPVNTQEAQMDMRGRLHSADSLKLSLSELRKTAIFISGLASUTENVD
RIETLSQAQEGTERLQIGEGAAAGSLKELEAHYSSVYVTPPDE"
/complement(42717..43520)
/gene="OSJNBa0083M16.8"
/complement(42717..43520)
/gene="OSJNBa0083M16.8"
/notes="hypothetical protein"
```



```
/codon_start=1
/protein_id="BAB64610.1"
/db_xref="GI:15528588"
/translation="MSMTSPRLKLFDFVSEETAEHEOYSAEPAPVCGNGGG
GGSDSSSTTTTATAGDGRREYCCRFANSLGQHONAKKERQOLKRA
QLOAAAAAGRVAGGAALIPRANPMVSAFAPPHLLGGGEAGPTSWYFSPRAA
VAGGQGFVHSGCVFARGAGPAVAAPAFVSTPATSAVAAPVYTTDEH
HGRIHASSAASLARYPGPMVAEPVATPVEDTCLGLDLQLSLAPAGL"
/gene="OSJNBa0083M16.9"
/complement(join(57300..57361,57501..57585,57688..57936))
/gene="OSJNBa0083M16.9"
/complement(join(57300..57361,57501..57585,57688..57936))
/gene="OSJNBa0083M16.9"
/feature="hypothetical protein"
/codon_start=1
/protein_id="BAB64611.1"
/db_xref="GI:15528589"
/translation="MSLAVATPASARSLPTTSSPECRRLLSAAPIRLRLR
RIAVASVPPAPPAEYTPWLIAGLNGPSKYHGRHNVGFEMVDRIARDEGITMT
IQSKSLGIAEITVVEAIDLVOKTTLTFV"
/complement(59958..63558)
/gene="OSJNBa0083M16.10"
/complement(59958..63558)
/gene="OSJNBa0083M16.10"
/feature="contains EST C26390(C12241)"

Query Match      37.6%; Score 67.6; DB 8; Length 179428;
Best Local Similarity 63.6%; Pred. No. 5.3e-10;
Matches 103; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 19 TATCTGATGACGTTGTTCAGCAACCTCTTTTACAAACCAATCGTTCTGAGTCATCTT 78
Db 129013 TACCTGTGCCAGTCGTCGAGGAGACGGCCTGGTACAAACAGATCTTCTCAGCCATGTG 129072
.QY 79 TTGCCGCGCATCTATCGGAACCTTACCTCATTTCTCCAGACATCGCTCCGAAATAC 138
Db 129073 GTCCCGHGGACTGGTGGCGGCCTCCGCCACCGCTCCAGTCTCGCTCCGCAACGGC 129132
QY 139 CTGCGCCGACACCTCATCTACTTCATCTCCGGTTTCTCTCGG 180
Db 129133 CTCGGCGACCTCATCTACTTCGCTCGGCTCTCTCTCGG 129174

RESULT 8
AF099969 Nicotiana tabacum mRNA linear PLN 19-MAY-1999
DEFINITION Nicotiana tabacum sterol-C5(6)-desaturase homolog mRNA, complete cds.
ACCESSION AF099969
VERSION AF099969.1 GI:4426626
KEYWORDS Nicotiana tabacum.
SOURCE Nicotiana tabacum
ORGANISM Nicotiana tabacum
REFERENCE 1 (bases 1 to 1148)
AUTHORS Russelstein,T., Schaller,H., Gachotte,D. and Benveniste,P.
TITLE Delta7-sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles
JOURNAL Plant Mo... Biol. 39 (5), 891-906 (1999)
MEDLINE 99273992
PUBMED 10344195
REFERENCE 2 (bases 1 to 1148)
AUTHORS Benveniste,P.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1998) Plant Molecular Biology Institute, C.N.R.S. 28 rue Goethe, Strasbourg 67083, France
FEATURES
source
1..1148
location/Qualifiers
/organism="Nicotiana tabacum"
/cultivar="xanthi SH6"
/db_xref="taxon:4097"
/cell_type="protoplast (calli)"
```

```
/tissue_type="leaf"
127..942
/function="introduces a double bond at C5 of
delta7-avenasterol resulting in
stigmasta-5,7,224(28)-trien-3beta-ol"
/codon_start=1
/product="sterol-C5(6)-desaturase homolog"
/protein_id="AAD20458.1"
/db_xref="GI:4426627"
/translation="MDDYLNLFIEETSFYNRVVLGTLPESWNGPLPHWFOGLRNLY
GGVLLYFISFLKCFYIRLRKNVYIPDAIPSRNMLQIGVAKMPFPCALPSSL
EYIMVNGWTKCFSDVGLWLSYLIYMAVLYIVFEGYIMHRELHDKLYKYLHAT
HHIYNKQNTLSPFAGLAFHPLDGLQAVPHVVALFLPEHFTIALFLFAITANI
HDCKHAKVWPVMGAGYHTIHHTYRHNYGHYTIWMDWFGTLRDPVEDEVKKM"
BASE COUNT 302 a 236 c 229 g 381 t
ORIGIN

Query Match      34.1%; Score 61.4; DB 8; Length 1148;
Best Local Similarity 60.5%; Pred. No. 3.9e-08;
Matches 101; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 14 ATGCTTATCTGATGCAGTTTGTTCAGCAACCTCTTTTACAAACCAATCGTTCTGAGTC 73
Db 131 ACGACTACCTGAATTTATTCATAGAGAGACATCGTTTACATCGGGTGGTTTATGATA 190
QY 74 ATCTTTTCCGCGGAATCTATGGAACCTTACCTCATTTCTCCAGACATCGTCCGAA 133
Db 191 CATCTTACCGGAATCATGTTGGTGGTCCACCTCTCTCATCTGTTTTCAGGGTGGTCCAGGA 250
QY 134 ATTACCTCGCGGGAACCTTACTATACCTCATCTCGGTTTCTCTCTG 180
Db 251 ATTACATTGGTGGGTGTACTTACTTATATATCTGTTTCTTCTTGG 297

RESULT 9
AX155063
LOCUS AX155063
DEFINITION Sequence 109 from Patent WO0138484.
ACCESSION AX155063
VERSION AX155063.1 GI:14536617
KEYWORDS Physcomitrella patens.
SOURCE Physcomitrella patens.
ORGANISM Physcomitrella patens
REFERENCE 1 (bases 1 to 1381)
AUTHORS Lerchl,J.; Renz,A., Ehrhardt,T., Reindl,A., Cirpus,P., Bischoff,F., Frank,M., Freund,A., Duwenig,E., Schmidt,R.M. and Reski,R.
TITLE Moss genes from Physcomitrella patens encoding proteins involved in the synthesis of polyunsaturated fatty acids and lipids
JOURNAL Patent: WO 0138484-A 109 31-MAY-2001; BASF Plant Science GmbH (DE)
FEATURES
source
1..1381
location/Qualifiers
/organism="Physcomitrella patens"
/db_xref="taxon:3218"
/feature="note"=pp004064012R"
/codon_start=1
/protein_id="CAC42667.1"
/db_xref="GI:14536618"
/translation="MASRGAVNMVCAIAIVLWMSLSLCSADVEYVNASFSVVG
GAKTKSGVVPANGSPYELALFVEETRWYNDLVGLPMLSPSVROSIPHTLQTLWRNY
AGMLLYFVSGGLWCLYVSWKGEHFPAGDIPAKEPIMLQIWTMKAMPYVTLGLPTLS
EYMLERGTWTKCFARIEDYVGLVAVLAVYVFGYIMHRELHDKLYKYLHAT
HHIYNKQNTLSPFAGLAFHPLDGLQAVPHVVALFLPEHFTHEVLLFCGCVTTNI
HDCIDGNVWGLMGAGFHTIHHTYRHNYGHYTIWMDWFGTLRDPYERKATFAHVKSS"
BASE COUNT 323 a 331 c 351 g 376 t
ORIGIN

Query Match      30.4%; Score 54.8; DB 6; Length 1381;
Best Local Similarity 59.7%; Pred. No. 4.1e-06;
Matches 92; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
```


TITLE
JOURNAL

Direct Submission
Submitted (20-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 174556)
Worley, K.C.

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18846184.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GIYE

Center clone name: CH230-103L10

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 114370 bases at least Q40

Consensus quality: 121991 bases at least Q30

Consensus quality: 127877 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1241: contig of 1241 bp in length
* 1242 1341: gap of unknown length
* 1342 2387: contig of 1046 bp in length
* 2388 2487: gap of unknown length
* 2488 3645: contig of 1158 bp in length
* 3646 3745: gap of unknown length
* 3746 5151: contig of 1406 bp in length
* 5152 5251: gap of unknown length
* 5252 6629: contig of 1378 bp in length
* 6630 6729: gap of unknown length
* 6730 8303: contig of 1574 bp in length
* 8304 8403: gap of unknown length
* 8404 9650: contig of 1247 bp in length
* 9651 9750: gap of unknown length
* 9751 10767: contig of 1017 bp in length
* 10768 10867: gap of unknown length
* 10868 12522: contig of 1655 bp in length
* 12523 12622: gap of unknown length
* 12623 14174: contig of 1552 bp in length
* 14175 14274: gap of unknown length
* 14275 15962: contig of 1688 bp in length
* 15963 16062: gap of unknown length
* 16063 17144: contig of 1082 bp in length
* 17145 17444: gap of unknown length
* 17446 18363: contig of 1119 bp in length
* 17447 18463: gap of unknown length
* 18464 19715: contig of 1252 bp in length
* 19716 19815: gap of unknown length
* 19816 21377: contig of 1562 bp in length
* 21378 21477: gap of unknown length
* 21478 22694: contig of 1217 bp in length
* 22695 22794: gap of unknown length
* 22795 24119: contig of 1325 bp in length
* 24120 24219: gap of unknown length
* 24220 25302: contig of 1083 bp in length
* 25303 25402: gap of unknown length

25403 26782: contig of 1380 bp in length
* 26783 26882: gap of unknown length
* 26883 28336: contig of 2054 bp in length
* 28337 29036: gap of unknown length
* 29037 30973: contig of 1837 bp in length
* 30974 32115: contig of 1442 bp in length
* 32116 32416: gap of unknown length
* 32417 34135: contig of 1620 bp in length
* 34136 34235: gap of unknown length
* 34236 35785: contig of 1530 bp in length
* 35786 35885: gap of unknown length
* 35886 37119: contig of 1233 bp in length
* 37120 37219: gap of unknown length
* 37220 3857: contig of 1639 bp in length
* 3858 38957: gap of unknown length
* 38958 40822: contig of 1665 bp in length
* 40823 40922: gap of unknown length
* 40924 42897: contig of 2175 bp in length
* 42898 42998: gap of unknown length
* 42999 4433: contig of 1436 bp in length
* 4434 44534: contig of 2341 bp in length
* 44535 46974: gap of unknown length
* 46975 48922: contig of 1948 bp in length
* 48923 49022: gap of unknown length
* 49024 51119: contig of 2097 bp in length
* 51120 51219: gap of unknown length
* 51220 53055: contig of 1836 bp in length
* 53056 53155: gap of unknown length
* 53156 55184: contig of 2029 bp in length
* 55185 55284: gap of unknown length
* 55285 56390: contig of 1106 bp in length
* 56391 56490: gap of unknown length
* 56491 58207: contig of 1717 bp in length
* 58208 58307: gap of unknown length
* 58308 61073: contig of 2766 bp in length
* 61074 61173: gap of unknown length
* 61174 62294: contig of 1121 bp in length
* 62295 62394: gap of unknown length
* 62395 63933: contig of 1539 bp in length
* 63934 64033: gap of unknown length
* 64034 66199: contig of 2166 bp in length
* 66200 66299: gap of unknown length
* 66300 68544: contig of 2245 bp in length
* 68545 68644: gap of unknown length
* 68645 71314: contig of 2670 bp in length
* 71315 71414: gap of unknown length
* 71415 73505: contig of 2091 bp in length
* 73506 73605: gap of unknown length
* 73606 76323: contig of 2718 bp in length
* 76324 76423: gap of unknown length
* 76424 7845: contig of 2322 bp in length
* 7846 78845: gap of unknown length
* 78846 80957: contig of 2112 bp in length
* 80958 81057: gap of unknown length
* 81058 82339: contig of 1282 bp in length
* 82340 82439: gap of unknown length
* 82440 84739: contig of 2300 bp in length
* 84740 84839: gap of unknown length
* 84840 87042: contig of 2203 bp in length
* 87043 87142: gap of unknown length
* 87143 90257: contig of 3115 bp in length
* 90258 90357: gap of unknown length
* 90358 92861: contig of 2504 bp in length
* 92862 92961: gap of unknown length
* 92962 95075: contig of 2114 bp in length
* 95076 95175: gap of unknown length
* 95176 99198: contig of 4023 bp in length
* 99199 99298: gap of unknown length
* 99299 102680: contig of 3382 bp in length

Query Match

19.9%; Score 35.8; DB 2; Length 174556;

```

Best Local Similarity 53.1%; Pred. No. 3;
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 38 ACGAACCTCTTTTACACCAACGATCTTCTGAGTCATCTTTGCGGCGAATCATGGG 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36620 ACAATGACCTCTTTACATATATTCATTCATGCACATCATGACTTGTATACCTATCTC 36679

QY 98 AACCCCTACCTGATTTTCCAGACATGCTCCGAAATACCTCCCGGAACCCACATAT 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36680 AACACTTACCAATTTCTCCATCTTACTCCGCGAGAACCTTCTCTCCCAATGGA 36739

QY 158 ACTTCATCTCCGTTTCTCTCG 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36740 ACAACTGTTTCTGCTGCTTTGG 36762

RESULT 12
AC007247/c
LOCUS Homo sapiens BAC clone RP11-305H21 116788 bp DNA linear PRI 14-OCT-2000
DEFINITION Homo sapiens BAC clone RP11-305H21 from Y, complete sequence.
ACCESSION AC007247
VERSION AC007247.5 GI:10801460
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 116788)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 116788)
Kalicki, J., Bemis, G. and Hawrysko, C.
The sequence of Homo sapiens BAC clone RP11-305H21
Unpublished
3 (bases 1 to 116788)
Waterston, R.H.
Direct Submission
Submitted (07-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 116788)
Waterston, R.
Direct Submission
Submitted (14-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 14, 2000 this sequence version replaced gi:9838137.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0305H21
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a

collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Center Institute (<http://bacpac.med.buffalo.edu>) VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-558K21, 200 bp overlap; the clone sequenced to the right is RP11-105L10, 200 bp overlap. Actual start of this clone is at base position 156866 of RP11-558K21; actual end is at base position 53998 of RP11-105L10.

The sequence RP11-305H21 from base position 114735 to 116434 contains a tandem repeat. The region was sized with PCR from clone DNA.

FEATURES	source	Location/Qualifiers
repeat_region		1..116788
repeat_region		/organism="Homo sapiens"
repeat_region		/db_xref="taxon:9606"
repeat_region		/chromosome="Y"
repeat_region		/map="Y"
repeat_region		/clone="RP11-305H21"
repeat_region		/clone_lib="RPC1-11"
repeat_region		3076..4236
repeat_region		/rpt_family="LI"
repeat_region		4244..4318
repeat_region		/rpt_family="LI"
repeat_region		4506..4807
repeat_region		/rpt_family="Alu"
repeat_region		5517..5697
repeat_region		/rpt_family="Alu"
repeat_region		5916..6328
repeat_region		/rpt_family="LI"
repeat_region		6602..6677
repeat_region		/rpt_family="MER1_type"
repeat_region		6714..6806
repeat_region		/rpt_family="MER1_type"
repeat_region		6834..6986
repeat_region		/rpt_family="MER1_type"
repeat_region		7529..7935
repeat_region		/rpt_family="MaLR"
repeat_region		7982..8316
repeat_region		/rpt_family="LI"
repeat_region		8995..9364
repeat_region		/rpt_family="LI"
repeat_region		9365..9654
repeat_region		/rpt_family="Alu"
repeat_region		9655..9670
repeat_region		/rpt_family="LI"
repeat_region		9666..9846
repeat_region		/rpt_family="LI"
repeat_region		10008..10584
repeat_region		/rpt_family="ERV1"
repeat_region		10600..10717
repeat_region		/rpt_family="LI"
repeat_region		10793..10950
repeat_region		/rpt_family="ERV1"
repeat_region		12680..13210
repeat_region		/rpt_family="MER1_type"
repeat_region		13241..13466
repeat_region		/rpt_family="MaLR"
repeat_region		14411..14515
repeat_region		/rpt_family="Alu"

repeat_region	15095..15403	/rpt_family="Alu"
repeat_region	15419..15845	/rpt_family="L1"
repeat_region	15904..16274	/rpt_family="Alu"
repeat_region	16310..16754	/rpt_family="MER1_type?"
repeat_region	16899..17207	/rpt_family="Alu"
repeat_region	17261..17490	/rpt_family="Alu"
repeat_region	17491..17579	/rpt_family="L1"
repeat_region	17849..18132	/rpt_family="Alu"
repeat_region	18369..18522	/rpt_family="Alu"
repeat_region	21067..21479	/rpt_family="L1"
repeat_region	21769..21844	/rpt_family="MR"
repeat_region	22148..22435	/rpt_family="Alu"
repeat_region	22441..22605	/rpt_family="L1"
repeat_region	22704..23406	/rpt_family="ERV1"
repeat_region	23439..23873	/rpt_family="ERV1"
repeat_region	23874..24331	/rpt_family="ERV1"
repeat_region	24411..25695	/rpt_family="ERV1"
repeat_region	26982..28044	/rpt_family="ERV1"
repeat_region	28045..28340	/rpt_family="Alu"
repeat_region	28341..28764	/rpt_family="ERV1"
repeat_region	28756..29011	/rpt_family="ERV1"
repeat_region	29005..29561	/rpt_family="ERV1"
repeat_region	29562..29987	/rpt_family="ERV1"
repeat_region	30054..30367	/rpt_family="Alu"
repeat_region	31201..31501	/rpt_family="Alu"
repeat_region	32016..32096	/rpt_family="ERV1"
repeat_region	32097..32389	/rpt_family="Alu"
repeat_region	32390..32928	/rpt_family="ERV1"
repeat_region	34746..35362	/rpt_family="L1"
repeat_region	35364..35467	/rpt_family="Alu"
repeat_region	35468..35976	/rpt_family="ERVK"
repeat_region	35977..36078	/rpt_family="Alu"
repeat_region	36125..36211	/rpt_family="Alu"
repeat_region	36212..36872	/rpt_family="L1"
repeat_region	37186..37479	/rpt_family="Alu"
repeat_region	38349..38412	/rpt_family="L1"

Query Match
Best Local Similarity 51.3%; Score 34.8; DB 9; Length 116788;
Matches 81; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 18 TTATCTGATCGAGTTGGTGACGAACCCTTCCTTATTACACCCCAANTCCTGTGCAGTCATCT 77
| | | | | | | | | | | | | | | |
Db 50771 TTTCTCCAGACATTTCTCTTAACAACAAGATAGTTCAAATCGGCCCTACTGAGCATCA 50712

QY 78 TTTGCGGGGAAATCTATGGGAACCCCCTACCTCATTTTCTCCAGACATGCTCCGAAAATA 137
| | | | | | | | | | | | | | | |
Db 50711 TTAGCAGTTGGTTTTATTTGCTATTTAAGTGCTTTTCCCCAACAGAAGGAGGACCCTIG 50652

QY 138 CCTCGCGGGAAACCTACTATATCTATCTCCGCTTCC 175
| | | | | | | | | | | | | | | |
Db 50651 CTAGCAAAAGTCTCTTCCAGAAGTCACATCCATTTTAC 50614

RESULT 13
LOCUS AC116209 linear HTG 13-JUL-2002
DEFINITION Rattus norvegicus clone CH230-16F3, *** SEQUENCING IN PROGRESS ***,
61 unordered pieces.
ACCESSION AC116209
VERSION AC116209.2 GI:21736254
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 186537)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.J., Bowie,S., Brivava,A., Brown,E., Brown,M., Bryant,N.P., Buhaty,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthevaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guervara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovac,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H., Lozato,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokoko,S., Ogwu,M., Okuwonu,G., Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Petersen,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savory,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taborski,P., Tamelisa,A., Tamelisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-More,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 186537)
Worley, K.C.
Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 186537)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:19718403.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWK
Center clone name: CH230-16F3
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143677 bases at least Q40
Consensus quality: 148892 bases at least Q30
Consensus quality: 153855 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1504: contig of 1503 bp in length
1604: gap of unknown length
2663: contig of 1059 bp in length
2763: gap of unknown length
4054: contig of 1291 bp in length
4154: gap of unknown length
4154: contig of 1345 bp in length
5499: gap of unknown length
5599: contig of 1271 bp in length
6870: gap of unknown length
6970: contig of 1173 bp in length
8143: gap of unknown length
8243: contig of 1188 bp in length
9431: gap of unknown length
9531: contig of 1348 bp in length
10879: gap of unknown length
10979: contig of 1060 bp in length
12039: gap of unknown length
12139: contig of 1776 bp in length
13915: gap of unknown length
14015: contig of 1355 bp in length
15370: gap of unknown length
15469: contig of 1485 bp in length
16954: gap of unknown length
17055: contig of 1191 bp in length
18245: gap of unknown length
18345: contig of 2493 bp in length
18346: gap of unknown length
20939: contig of 1496 bp in length
22434: gap of unknown length
22435: contig of 1548 bp in length
22535: gap of unknown length
24083: contig of 1685 bp in length
24182: gap of unknown length
25867: contig of 1685 bp in length
25868: gap of unknown length

25968: contig of 2159 bp in length
28127: gap of unknown length
28226: contig of 1590 bp in length
28227: gap of unknown length
29817: contig of 2722 bp in length
32639: gap of unknown length
32738: contig of 2274 bp in length
35012: gap of unknown length
35112: contig of 1353 bp in length
35113: gap of unknown length
36466: contig of 2816 bp in length
36566: gap of unknown length
39382: contig of 2483 bp in length
41964: gap of unknown length
42064: contig of 1645 bp in length
43709: gap of unknown length
43809: contig of 2811 bp in length
46620: gap of unknown length
46720: contig of 1758 bp in length
48478: gap of unknown length
48579: contig of 2466 bp in length
51044: gap of unknown length
51144: contig of 1562 bp in length
52706: gap of unknown length
52806: contig of 2427 bp in length
55233: gap of unknown length
55334: contig of 1786 bp in length
57119: gap of unknown length
57219: contig of 2682 bp in length
59901: gap of unknown length
60001: contig of 2112 bp in length
62113: gap of unknown length
62214: contig of 2034 bp in length
64247: gap of unknown length
64347: contig of 1688 bp in length
66015: gap of unknown length
66115: contig of 3288 bp in length
69403: gap of unknown length
69504: contig of 2535 bp in length
72038: gap of unknown length
72138: contig of 3696 bp in length
75834: gap of unknown length
75934: contig of 3070 bp in length
79004: gap of unknown length
79104: contig of 3136 bp in length
82240: gap of unknown length
82340: contig of 2259 bp in length
84599: gap of unknown length
84600: contig of 3493 bp in length
88192: gap of unknown length
88292: contig of 3181 bp in length
88293: gap of unknown length
91474: contig of 2693 bp in length
91574: gap of unknown length
94267: gap of unknown length
94367: contig of 3737 bp in length
98104: gap of unknown length
98204: contig of 3066 bp in length
101269: gap of unknown length
101369: contig of 3377 bp in length
104746: gap of unknown length
104846: contig of 3191 bp in length
108037: gap of unknown length
108137: contig of 3830 bp in length
111967: gap of unknown length
112067: contig of 3077 bp in length
115144: gap of unknown length
115245: contig of 2693 bp in length
117937: gap of unknown length
118038: contig of 3547 bp in length
121584: gap of unknown length
121585: contig of 5578 bp in length
127262: gap of unknown length
127363: contig of 3741 bp in length


```
gene
CDS
/db_xref="taxon:71421"
165..1403
/gene="H11348"
165..1403
/gene="H11348"
/notes="similar to SP:P29745 GB:U00096 PID:1651556
PID:1787372 percent identity: 54.27; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptidase T (pepT)"
/protein_id="AAC22995.1"
/db_xref="GI:1574810"
/translation="MISQIDKTELLERLEFHVYFHTQSKPNKAKHSPSSVGOMKLAQOL
QKELTQGLLENVEVSKYAVVTFPLANDPNLTIGLVVAHLDTSPOCSGKNVRPVEIV
EYRGDIALGEEIEISPVYSFMOKLVGQTLIVDTGTLTGADNKAQIAETMTALSI
LOKENIPICNIRVATPDEEIGLGHYFPEMEKFSQWYATIDGGVGELEVENFNAAT
AKVRFGRNHFYAGKGMNALTACCEQQVFPVDEPEKTDGKVGFGYHLEDFSGDI
EVEITYLIRDEQONFAQKAFIKNOVEKFNKGLKKPIELEIQDSYQNMVDVKN
VQSIELADRAMKAVGIRPNHKPIRGGTGAPLASKGLACPNIFTGGYNFHSKHELVS
LQGMENYVQVITMLKCKDL"
complement(1967..2449)
/gene="H11349"
complement(1967..2449)
/gene="H11349"
/notes="similar to PID:1652428 percent identity: 31.37;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22996.1"
/db_xref="GI:1574811"
/translation="MSKTSIGLKVQSAELADKLNELLATYQVFTNVRGVHWNKVG
NFFALHAKFEETYNLVARVDEVAERILTVTPNNAYSOVLKSRKEDIASVQAE
CLSGTLQGLKTLIDQOREILAFANNANDEGTASQMSDIKEQEKLVWFPQACQTCN
2723..3601
/gene="H11350"
2723..3601
/gene="H11350"
/notes="similar to GB:X63144 GB:U15410 GB:X16419 PID:41090
PID:560012 percent identity: 51.18; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="cytidine deaminase (cdd)"
/protein_id="AAC22997.1"
/db_xref="GI:1574812"
/translation="MQDLIKRTLPODDALNQAIYVNDLRSONWAGELNYSQVOQLCHNF
EUTLUKLAHLLPLAASYSHTAISFNVAIGAIGQDDFYFGANQEFANSAIQOTIHA
EQSAISHAWLRNERRISDMVNYTPCGHCRQPMNELHGAEKISIHLPSONNPLSHYL
PDAGFKDLDTAAHLLAEENHDLVADHODLVNQAILAANQSHCPYSNSPHGIALIFK
NSDVVTGRVAENAFNPSPALQALNFAYLNDKKLSLSDIERIVMAEKALKLSHKTMAE
TLLSLTSVELEYYSL"
complement(3644..4609)
/gene="H11351"
complement(3644..4609)
/gene="H11351"
/notes="similar to GB:U00096 PID:1788178 percent identity:
65.22; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22998.1"
/db_xref="GI:1574813"
/translation="MIDRPFYQIATTNLSDWLETLPCOLKEWETQTHGYAKWSKI
VDFNLNHADEIDLKSAKSDRTSLDGGCGSYHMMRWGEGAKMVGIDPFLFICQF
EWRSDFKWDRVLPHLSPLOGRITLDVGGSGYHMMRWGEGAKMVGIDPFLFICQF
EAVRLLNDRANILPIGIEQMPPLAADFVTFMSGVLYHRRKSLDHLHSOLKNQLVKG
GELVLETVDDGDIINTVLVPADYAKMKNVYFIPSVATLINNLEKVGFTNVRCDVAT
TFLERQKTDWLENESLIDFLDPNDHSKTIIEGYQAPKRAVILANK"
complement(4606..6120)
/gene="H11352"

CDS
complement(4606..6120)
/gene="H11352"
/notes="similar to GB:D32069 PID:509694 PID:913690 percent
identity: 64.98; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="sodium/proline symporter (proline permease)
(putP)"
/protein_id="AAC22999.1"
/db_xref="GI:1574814"
/translation="MFGDPDSLITFTIYIFGMLLGLVLAAYVTNNLSYVILGGRRLGS
FVTSAGASOMSGWLLMGLPGAVYLSGLVEGIAITGLTYGAYFNMLLVAGRLRYTE
LNNALTIPEYFHNFGSHKLLKLVSATIILVFTIYCASCVAGAKLQFIHFSVEY
STALWYGAATAYTFIGGLAVSWDTIOATLMIFALITPVFVLLSFADTAQFSAY
LEQAAVYNKQFTDLFTSTPLGLLSLAAMGLGYGQPHILAREFNAADSVKSLIKARR
ISMWVWVLCLAGAIGLGAIPYFFANPAIAGTVNREPEQVFIELAKLLFNPIWAGIL
LSAILAAVAGTSLAQLLSSISITEDFYKGFIRPNASEKELVNLGRIMVLYIAALAIW
TAQDENSVKLVLEPAGFSAFGPVVLFSLFWKRMFTSSGAMGLMGLVAVTVFWAKE
VVPADTDKFKVYEMIPGFAFASLAIIVISLSNKKPEQDILNTFKAERAKYKAEK"
6236..7711
/gene="H11353"
6236..7711
/gene="H11353"
/notes="similar to SP:P25537 PID:48826 PID:606187 GB:U00096
PID:1789645 percent identity: 70.96; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cytoplasmic axial filament protein (cafa)"
/protein_id="AAC23000.1"
/db_xref="GI:1574815"
/translation="MDAVELLMNVTPTNETRIALVETGLMREVHIERQAKRGIVGNIYK
GRVTLPMQOSAFVDIGLEAAFLHAAIDIVSHTECVDENEKQKFKVKSISLVREGO
DIVQVWKEPCTGKARLTDTITLSRHLVFPENSHVGYQSRIESEERARLKAIVE
PFCDELGIINTATEGASEEELRODAEFLKRLKRVLRKRSKYTKTKYGEIPALPO
RILRDFIGNLLEKIRIDSKLFCGEKFTDEPMPELSDKLVLSQNPFIYGVYVNA
IQTLALGRVNLKSGYLIETQTEAMTTIDINTGAFVGHRLNEETIFNTIEATKAIAH
ELQRLNGLGIIIDIDMOTDEHRNRVLQSLCDALSCKDRMKTNVNGFTQLGLVEMTRK
RTRESLHVLDCECTCHGRGVKTVETVCYEMREIRVYHLSFSESEQFVVYASPAVS
EYLINESHGLLPEVEMFGRKRVKTEQFYNQEQFDVVVM"
8131..9804
/gene="H11354"
8131..9804
/gene="H11354"
/notes="similar to GB:V01575 SP:P00962 GB:J01617 GB:M0187
PID:146168 percent identity: 76.28; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutaminyl-tRNA synthetase (glns)"
/protein_id="AAC23001.1"
/db_xref="GI:1574816"
/translation="NMSHTETSLGAENTRTHNFITQIIDDLASCKGSHVHTFPFPEP
NGYLHGHAKSICLNFGLAKETQGCNLRFDFTNPKVEDVEYVDSIKADVEMLGFKEW
GEPYADYFDALYGYAVLEIKKGLAYVDELSPEDEMREYGTLPFGKNSPYRDRITIE
ENLALFERKNGEFAEGKASRAKIDMASPFMREPVYIRKFSHHQTDGKWCIYP
MYDFTHCISDAIERHTSCTLEFQDNRLYDWLENISIERLPHQVEFSRLNLEGT
LTSKRLKLVNDEIVDGNQDPMPTISGLRRGYTPASLREFCRRIGVTKODNVVEYK
SALEACIREDLNENAPRAMVIDPVRVVTFENFESEAVLTAPNHPRLGEGRLPFTK
ELYIDRADFREANKQYKRLVGLKGVRLNAYIVKAEVRKDAKEGIEITICTYDPT
LGKNPADGRKGVKGIHWYSAVNNHHPAEFLYDLRFTVPNPGAEDDIESVLPNSLVIK
QGFEQSLANAEAEKGYQFEREGYFCADSKDSRPEHLVFNLTVSLKEGF"
9870..10334
/gene="H11355"
9870..10334
/gene="H11355"

Query Match 18.1%; Score 32.6; DB 1; Length 10393;
Best Local Similarity 54.6%; Pred. No. 28;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 59 GAATCGTCTGAGTCATCTTTTGGCGGCGGAATCTATGGGAACCCCTACCTCATTTCTCC 118
```

```

      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    8136 GAGCCAATACAGAGACATCTTTGGGTGCAGAAATACCCGCACTCATAATTTATTACCCA 8195
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy    119 AGACATCGCTCCGAAATTACCTCCGCGAACCCCTACTATCTTCATCTCGGTTTCCTC 177
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    8196 AATTATTGATGAAGATTAGCTTCGGGTAACATAAAAGGTTTCATACCCGTTTTCGCG 8254

```

Search completed: June 6, 2003, 15:13:59.
 Job time : 529.302 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nuclei: search, using sw model

Run on: June 6, 2003, 12:27:48 ; Search time 467.887 Seconds
(without alignments)

6230.541 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322

Perfect score: 180

Sequence: 1 a1agggcggaatagttta.....tcattctcgggttctctg 180

Scoring table: IDENTITY_NUC

Gap 10.0, Gapext 1.0

Searched: 161:4066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othe:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.4	49.0	551	10	BE471892
2	88.4	49.0	597	12	BG098248
3	88.4	49.0	672	12	BG600842
4	87.8	48.0	572	10	BE471514
5	87.8	48.0	608	14	BQ118369
6	86.8	48.0	513	10	AW429091

7	85.2	47.3	753	13	BI930355
8	83.8	46.6	621	12	BG123510
9	83.6	46.4	567	10	AW429266
10	76.4	42.4	486	14	BQ028052
11	76	42.2	535	10	AW596303
12	76	42.2	583	12	BF425767
13	75	41.7	722	12	BG598293
14	71.2	39.6	444	13	BI470419
15	70.8	39.3	535	13	BW522429
16	69.6	38.7	454	10	AW624414
17	69.6	38.7	458	9	AU163003
18	68.8	38.2	592	10	AW219366
19	67.6	37.6	359	9	AU068067
20	67.6	37.6	458	9	AU070031
21	58	32.2	405	12	BG487492
22	58	32.2	541	12	BG322635
23	58	32.2	562	10	BE125051
24	58	32.2	570	13	BG933493
25	58	32.2	587	12	BF586850
26	58	32.2	667	10	BE357415
27	55.4	30.8	530	13	BJ200021
28	54.8	30.4	604	13	BJ195304
29	54.8	30.4	607	13	BJ202813
30	54.8	30.4	645	13	BJ157215
31	54.6	30.3	804	13	BI308419
32	53.2	29.6	252	17	BH215130
33	53.2	29.6	605	13	BI417093
34	53.2	29.6	605	13	BI674125
35	52.8	29.3	716	10	AW75288
36	52	28.9	825	10	BE643026
37	50.4	28.0	586	14	BW903391
38	48	26.7	517	13	BI188385
39	44.6	24.8	483	13	BJ199974
40	44.6	24.8	492	13	BJ197093
41	44.6	24.8	517	14	BQ703739
42	42	23.3	306	12	BF483844
43	42	23.3	409	9	AL819228
44	42	23.3	658	14	BQ246840
45	40.6	22.6	325	12	BG265811

ALIGNMENTS

RESULT 1:
BE471892 LOCUS BE471892 551 bp mRNA linear EST 28-JUL-2000
DEFINITION EST416745 potato stolon, Cornell University Solanum tuberosum cDNA clone cSTA30N16, mRNA sequence.
ACCESSION BE471892
VERSION BE471892.1 GI:9562383
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 551)
AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R., Holt, I.E., Liang, F., Hansen, T.S., Utterback, T., Bowman, C.L., Doan, B., Bougri, O., Buehl, C.R., Ronning, C.M., Tanksley, S.D. and Baker, B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
FEATURES
Location/Qualifiers
1..551
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"

BI930355 EST550244
BG123510 EST469146
AW429266 EST306722
BQ028052 OHG9K09.Y
AW596303 sJ01f11.Y
BF425767 sr42f04.Y
BG598293 EST496971
BI470419 sah91f12.
BW522429 sak98h05.
AW624414 EST322359
AU163003 AU163003
AW219366 EST301848
AU068067 AU068067
AU070031 AU070031
BG487492 EML_65.G0
BG322635 EML_41.D0
BE125051 DGI_15.H1
BG933493 WSI_4.B01
BF586850 FMI_30.A1
BE357415 DGI_15.H1
BJ200021 BJ200021
BJ195304 BJ195304
BJ202813 BJ202813
BJ157215 BJ157215
BI308419 EST529829
BH215130 1006025E0
BI417093 949053A09
BI674125 949053A09
AW75288 EST334353
BE643026 Cr12_7.L0
BM903391 NXR_042
BJ188385 BJ188385
BJ199974 BJ199974
BJ197093 BJ197093
BQ703739 946105C08
BF483844 WHE2307.D
AL819228 AL819228
BQ246840 TAE15004G
BG265811 1000081H0

BASE COUNT ORIGIN	155 a	138 c	125 g	179 t	frozen in liquid nitrogen immediately upon removal from tubers."
Query Match	49.1%	Score 88.4;	DB 12;	Length 597;	
Best Local Similarity	68.5%;	Pred. No. 1.6e-17;			
Matches 122;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps 0;	
Qy	3	GGCGGGCGGAATAAGCTTATCTGATGCAGTTGTTGGTGGAGAAACCTCTCTTTTACAAACCGAAT	62		
Dd	70	GTCGGAGATGGAGATTACTTGAAGCTGTCGTGGAGAGACATCGTTTACAAACCGTAT	129		
Qy	63	CGTTCTGAGTCATCTTTTGGCGGGGAATCTATGTGGAACCCCTTACCTCATTTTCTCCAGAC	122		
Dd	130	GTTTTCGGTGACATTCTTTCGCGGAATCATGCTGGGGACCACCTTCTCATATGCTTCAAGG	189		
Qy	123	ATGCTCCGGAATTAACCTCGCGGGAACCCCTACTACTTCTATCCCGGTTTCTCTCGG	180		
Dd	190	ATGGCTCCGTAACCTACATGTCGGGTGTTTACTTACTTCTATCTCCGGTTTCTCTCGG	247		
RESULT 3					
RG600842					
LOCUS	EST505737	cSTS Solanum tuberosum	672 bp	mRNA	linear EST 12-APR-2001
DEFINITION	mRNA sequence.				
ACCESSION	RG600842				
VERSION	RG600842.1	GI:13617978			
KEYWORDS	EST.				
SOURCE	potato.				
ORGANISM	Solanum tuberosum				
REFERENCE	van der Hoeven R., Bezzerides J., Sun H., Cho J., Chiemiango A.,				
AUTHORS	Bougri, O., Buell, C. R., Ronning, C., Tanksley, S. and Baker, B.				
TITLE	Generations of ESTs from sprouting potato eyes				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Cathy Ronning				
	The Institute for Genomic Research				
	For clone info: please contact Research Genetics, Libraries				
	Division tel 1-800-711-6195, email cdna@resgen.com				
	Seq primer: M13P-R.				
FEATURES	Location/Qualifiers				
source	1..672				
	/organism="Solanum tuberosum"				
	/cultivar="Kennebec"				
	/db_xref="taxon:4113"				
	/clone="cSTS30013"				
	/clone_lib="cSTS"				
	/tissue_type="sprouting eyes from tubers"				
	/dev_stage="12-14 weeks post harvest"				
	/lab_host="SOLR"				
	/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."				
BASE COUNT	167 a	163 c	137 g	205 t	
ORIGIN					
Query Match	49.1%	Score 88.4;	DB 12;	Length 672;	
Best Local Similarity	68.5%;	Pred. No. 1.7e-17;			
Matches 122;	Conservative 0;	Mismatches 56;	Indels 0;	Gaps 0;	
Qy	3	GGCGGGCGGAATAAGCTTATCTGATGCAGTTGTTGGTGGAGAAACCTCTCTTTTACAAACCGAAT	62		
Dd	111	GTCGGAGATGGAGATTACTTGAAGCTGTCGTGGAGAGACATCGTTTACAAACCGTAT	170		
Qy	63	CGTTCTGAGTCATCTTTTGGCGGGGAATCTATGTGGAACCCCTTACCTCATTTTCTCCAGAC	122		

```

171 GGTGTTTGGTACATCTTGGCGGAATCATGTTGGGACACATTCCTCATATGCTTCAAGG 230
LOCUS
DEFINITION
123 ATGGCTCGGAATTAACCTCGCGGAACCTACTATACATCTCTCGGTTTCTCTGG 180
Db
231 ATGGCTCGTAACACTACATGGCGGTGTTTACTTTACTTCACTCCGGTTTCTCTGG 288

RESULT 4
BE471514
LOCUS
DEFINITION
EST416367 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA29120, mRNA sequence.
BE471514
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 572)
AUTHORS
van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Boufri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE
Generation of ESTs from potato swelling stolons
JOURNAL
Unpublished (1999)
CONTACT: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cna@resgen.com
5 prime sequence.

FEATURES
Location/Qualifiers
1..572
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTA29120"
/tissue_type="potato stolon, Cornell University"
/stolons="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT 135 a 137 c 122 g 178 t
ORIGIN
Query Match 48.8%; Score 87.8; DB 10; Length 572;
Best Local Similarity 68.2%; Pred. No. 2.5e-17;
Matches 122; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 2 TGGCGCGGATAATCGTTATCGATGCGAGTTGTTGGACGAACCTCTTTTACACCGAA 61
Db
113 TGTGCGAGATGAGGACTTACTTGAAGCTGTTCGTGGAGAGACTTCGTTTTACACCGTC 172
QY 62 TCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACCTCATTTTCTCCAGA 121
Db
173 TGGTTTGGGTATACATCTTTCGCGGGAATCATGTGGGGACCATCTCTCATATGCTTCAAG 232
QY 122 CATGCTCGGAATTAACCTCGCGGAACCTACTATATCTATCTCCGGTTTCTCTGG 180
Db
233 GATGCGTCGGTAACACTATGCGGTGTTTACTTACTTCACTATCTCCGGTTTCTCTGG 291
```

```

RESULT 5
BQ118369
LOCUS
DEFINITION
EST603945 mixed potato tissues Solanum tuberosum cDNA clone STMEA40
5' end, mRNA sequence.
BQ118369
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 608)
AUTHORS
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karanycheva,S.A.
TITLE
Generation of a set of potato cDNA clones for microarray analyses
JOURNAL
Unpublished (2002)
COMMENT
On Apr 17, 2002 this sequence version replaced gi:20170331.
CONTACT: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@tigr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cna@resgen.com
Seq primer: T3.

FEATURES
Location/Qualifiers
1..608
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binjfe"
/db_xref="taxon:4113"
/clone="STMEA40"
/clone_lib="mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."
BASE COUNT 155 a 144 c 125 g 184 t
ORIGIN
Query Match 48.8%; Score 87.8; DB 14; Length 608;
Best Local Similarity 68.2%; Pred. No. 2.5e-17;
Matches 122; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 2 TGGCGCGGATAATGCTTATCTGATGCGAGTTTGTGACGAACCTCTTTTACACCGAA 61
Db
86 TGTGCGAGATGAGGACTTACTTGAAGCTGTTCGTGGAGAGACTTCGTTTACACCGTC 145
QY 62 TCGTTCTGAGTCATCTTTTGGCGGGAATCTATGGGAACCTTACCTCATTTTCTCCAGA 121
Db
146 TGGTTTGGGTATACATCTTTCGCGGGAATCATGTGGGGACCATCTCTCATATGCTTCAAG 205
QY 122 CATGCTCGGAATTAACCTCGCGGAACCTACTATATCTATCTCCGGTTTCTCTGG 180
Db
206 GATGCGTCGGTAACACTATGCGGTGTTTACTTACTTCACTATCTCCGGTTTCTCTGG 264

RESULT 6
AW429091
LOCUS
DEFINITION
EST306463 tomato flower buds 0-3 mm, Cornell University
Lycopersicon esculentum cDNA clone cTOA2A22 5', mRNA sequence.
AW429091
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```


QY	3	GGCGCGGATAATGCTTATCTGATGAGTTTGTGTGACGAAACCTCTTTTACAAACGGAAT	62
Db	63	GGCGGAGATGGAGGATTACTTGAAGCTATTCTGTGGAGGAGACATCGTTTTTACAAACGGGT	122
QY	63	CGTTCCTGAGTCATCTTTTCCGCGGGAATCTATGCGGGAACCTTACCTCATTTTCTCCAGAC	122
Db	123	GGTTTTGGGTACATCTTCCCGGAATCATGGTGGGACCACCTTCCTGATGATGTTCAAGG	182
QY	123	ATGGCTCCGAAATTAATCTCGCGGAACCTTACTATCTATCTATCTCGGGTTTCTCTCGG	180
Db	183	ATGGCTCCGTAACATACATGGCGGGGTTTACTTTACTTCACTATCTCGGGTTTCTCTCGG	240
<p>RESULT 10 BU028052/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM</p>			
<p>BU028052 486 bp mRNA linear EST 23-AUG-2002 QHG9K09 yq ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus CDNA clone QHG9K09, mRNA sequence. BU028052 BU028052.1 GI:22463572 EST. common sunflower. Helianthus annuus</p>			
<p>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus 1. (bases 1 to 486)</p>			
<p>REFERENCE AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K. Lettuce and Sunflower ESTs from the Compositae Genome Project http://comgenomics.ucdavis.edu/ Unpublished (2002)</p>			
<p>JOURNAL COMMENT Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Asmunsdon Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu] singleton, see http://cgdb.ucdavis.edu/ for details. Plate: QHG9 row: K column: 09.</p>			
<p>FEATURES source 1..486 /organism="Helianthus annuus" /cultivar="RHA280" /db_xref="taxon:4232" /clone="QHG9K09" /clone_lib="QH_EFGHJ sunflower RHA280" /lab_host="E.coli" /note="Vector: pBRCNDSfiAB: The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/ TAG_LIB-QH_EFGHJ sunflower RHA280 TAG_TISSUE=developing kernel TAG_SEQ="GCCATCGGG"</p>			
<p>BASE COUNT 127 a 63 c 141 g 155 t ORIGIN</p>			
<p>Query Match 42.4%; Score 76.4; DB 14; Length 486; Best Local Similarity 69.3%; Pred. No. 9 9e-14; Matches 104; Conservative 0; Mismatches 46; Indels 0; Gaps 0;</p>			
QY	31	TTTGTTGACGAAACCTCTTTTACAAACCGAATCGTTCTGATGATCTTTTTCGCGGGAAT	90

```

Db      310 TTCGTCACAAAGAACTCTTTTCACACCAAAATAGTCTCGGAACCTTTACTGCCGGAACAA 251
QY      91 CTATGGAAACCTTACCTCATTTTCAGACATGGCTCCGAAATATACCTCCGCGAACC 150
        || || || || || || || || || || || || || || || || || || || || || ||
Db      250 CTCTGAGCCCACTCCGCCACCCCTTCAAACTGGCTCCGTAATACATTCGCGATCC 191
        || || || || || || || || || || || || || || || || || || || || || ||
QY      151 CTACTATATCATCTCCGGTTTCCTCTGG 180
        || || || || || || || || || || || || || || || || || || || || || ||
Db      190 CTGTTTACCTCATCTCTGGCTTCCTTTGG 161

RESULT 11
AW596303 535 bp mRNA linear EST 03-DEC-2001
sj01f11.y1 Gm-c1032 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1032-646 5' similar to TR:Q9ZT29 Q9ZT29 STEROL-C5(6)
)-DESATURASE. ;, mRNA sequence.
AW596303 1 GI:7283699
EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 535)
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1224 Std Error: 0.00
High quality sequence stop: 380.
Location/Qualifiers
1..535
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1032-646"
/clone_lib="Gm-c1032"
/tissue_type="Cotyledons of 8-day-old 'Williams'
seedlings"
/lab_host="DH10B"
/notes="Vector: pBluescript II XR; Site:1: EcoRI; Site:2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 8-day-old 'Williams' seedlings which
were propagated on paper towels with distilled water for
3 days (etiolated), then greenhouse grown for 5 days in
potting soil. The cotyledons were flash-frozen in liquid
nitrogen. StrataGene's cDNA synthesis kit (catalog number
200401) was used to synthesize the cDNA. First-strand
synthesis was performed with 5-methyl dCTP, hence the
ligated cDNA was hemimethylated. A modification of
StrataGene's first-strand synthesis primer was used. An
'anchor' nucleotide (V-A, C, or G) was added to the 3' end
of the primer [GAGAGAGAGAGAGAGAACTGTCAG(T)18V] to
anchor the primer at the 5' end of the poly(A) tract.
After second strand synthesis, the cDNA ends were filled
in with cloned pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The cDNA was then

```

```

precipitated and redissolved in sterile, RNase-,
DNase-free water. The xhoI site within the first-strand
synthesis primer was then restricted by digestion with
XhoI from Promega (400/ul); all xhoI sites in the cDNA
would be protected by their hemimethylated status. The
cDNA constructs were size-fractionated with a 500 bp
cutoff, using GibcoBRL Life Technologies' cDNA Size
Fractionation column. The column eluent was then
precipitated, redissolved, and ligated into StrataGene's
pBluescript II XR Predigested vector (pBluescript II SK(+))
vector that has been digested with EcoRI and XhoI, and
phosphorylated by StrataGene). 100% of the white and 87.5%
of the blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=28 and 8
respectively). This library was constructed by Dr. Paul
Keim and Dr. Virginia Corvelli."
BASE COUNT 121 a 138 c 110 g 163 t 3 others
ORIGIN

Query Match 42.2%; Score 76; DB 10; Length 535;
Best Local Similarity 65.1%; Pred. No. 1.4e-13;
Matches 112; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY      9 GGATATGCTTATCTGATGCACTTTGTTCAGCAACCTCTTTTACAAACCGAATCGTTCT 68
        || || || || || || || || || || || || || || || || || || || || || ||
Db      1 GAACAGCGAATATTCGAGGCTGTCTCGAGGACACGAGACTTGTACAATCGCATTTGT 60
        || || || || || || || || || || || || || || || || || || || || || ||
QY      69 GAGTCATCTTTTTCGGCGGGAATCTATGGAACCTTACCTCATTTTCTCCAGACATGGCT 128
        || || || || || || || || || || || || || || || || || || || || || ||
Db      61 GGAACCCCTCTTTCGCGCACTTCGTGTGGGGCCCACTCCGCACTTCTCCAGACATGGCT 120
        || || || || || || || || || || || || || || || || || || || || || ||
QY      129 CGAAATTAACCTCGCGGGAACCTACTATCTATCTATCTCCGGTTTCTCTCTGG 180
        || || || || || || || || || || || || || || || || || || || || || ||
Db      121 TCGCAATTACCTCGCGGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGG 172
        || || || || || || || || || || || || || || || || || || || || || ||

RESULT 12
BF425767 583 bp mRNA linear EST 06-DEC-2001
sr42f04.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1051-728 5' similar to TR:Q9ZT29 Q9ZT29 STEROL-C5(6)
)-DESATURASE. ;, mRNA sequence.
BF425767 1 GI:11413756
EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 583)
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Corvelli,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 393.
Location/Qualifiers
1..583

```

```
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-728"
/clone_lib="Gm-cl051"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
BASE COUNT      131 a 149 c 120 g 183 t
ORIGIN
Query Match      42.2%; Score 76; DB 12; Length 583;
Best Local Similarity 65.1%; Pred. No. 1.4e-13;
Matches 112; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 9 GGATAAGCTTATCTGATGCGAGTTGTTGAGCAACCTCTCTTTTACACCGAATCGTTCT 68
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GAACAG:GAATATTCGAGGCTGTTGTCGAGGACGACGACGTGTACATCGCATGTGTT 60
QY 69 GAGTCATCTTTGGGGGCAATCTATGAGGACCTTACTCTCATTTTCTCCAGACATGGCT 128
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GGGAACTCTCTTGGCGCACTTCGTGTGGGGCCCACTCCGCGACCTCTCCAGACGTGGCT 120
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 CGGAATTTACTCGCGCAACCTACTATCTATCTCCGCTTCCTCTCTCTG 180
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TCGCAATTTACTCGCGCGGCTCTCTCTACTCTCTCCGCGCTCTCTGTGG 172
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 13
LOCUS      BG598293
DEFINITION BG598293 722 bp mRNA linear EST 12-APR-2001
            Solanum tuberosum cDNA clone cSTS20L7 5' sequence,
            mRNA sequence.
ACCESSION  BG598293
VERSION    BG598293.1 GI:13616433
KEYWORDS   EST.
SOURCE     potato.
ORGANISM   Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
            van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
            Bougri,C., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
            Generations of ESTs from sprouting potato eyes
            Unpublished (2000)
            Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cduatresgen.com
            Seq primer: M13P-R.
            Location/Qualifiers
              1..722
                /organism="Solanum tuberosum"
                /cultivar="Kennebec"
                /db_xref="taxon:4113"
                /clone_lib="cSTS"
                /dev_stage="12-14 weeks post harvest"
                /lab_host="SOLR"
            /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
            taken from tubers. The tubers were incubated at 26C in the
            dark for 2-3 weeks prior to sprouting. The eyes were
            frozen in liquid nitrogen immediately upon removal from
            tubers."
BASE COUNT      191 a 165 c 148 g 218 t
ORIGIN
Query Match      41.7%; Score 75; DB 12; Length 722;
Best Local Similarity 71.1%; Pred. No. 3.1e-13;
Matches 113; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 23 TGATCGAGTTGTTGACGAA-ACCTCTTTTACACCGAATCGTTCTAGTCATCTTTTG 81
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 TGAAGCTCTTCGTGAGGAAGACATCGTTTACAAACCGTATGTTTGGGTACATCTTG 63
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 82 CCGGGAATCATATGGAACCCCTTACCTCTCATTTTCTCCAGACATGCTCCGAATACCTC 141
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 64 CCGGAATCATGTGGGACCACTTCTCATATGCTTCAAGGATGCTCCGTAACTACATT 123
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 142 GCCGAACCCCTACTATCTTCTATCTCCGCTTCTCTCTGG 180
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 124 GCGGCTGTTTACTTACTTCTATCTCCGCTTCTCTCTGG 162
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 14
LOCUS      BI470419
DEFINITION BI470419 444 bp mRNA linear EST 29-NOV-2001
            sah9lf12.v1 Gm-cl050 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-cl050-3360 5' similar to TR:Q9SWU4 Q9SWU4 DELTA7 STEROL C-5
            DESATURASE. [2] TR:Q99208 ;, mRNA sequence.
ACCESSION  BI470419
VERSION    BI470419.1 GI:15286528
KEYWORDS   EST.
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
            Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna
            A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            R., Waterston,R. and Wilson,R.
            Public Soybean EST Project
            Unpublished (1999)
            Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Putative full length read
            vector to vector length is 523 This clone is available through:
            Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
            35801 For further information call: (800)-533-4363 or contact via
            email: cduatresgen.com
            High quality sequence stop: 422.
            Location/Qualifiers
              1..444
                /organism="Glycine max"
                /db_xref="taxon:3847"
                /clone="GENOME SYSTEMS CLONE ID: Gm-cl050-3360"
                /clone_lib="Gm-cl050"
                /tissue_type="leaf tissue at various developmental stages
                of greenhouse grown plants"
                /dev_stage="3 week old"
                /lab_host="DH10B"
            /note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
            XhoI; The Clark NIL was constructed and seed was provided
```


Result No.	Score	Queri	Match	Length	DB	ID	Description
1	180	100.0		846	9	US-09-938-842A-421	Sequence 421, Appl
2	180	100.0		1889	10	US-09-775-879-20	Sequence 20, Appl
3	125.2	69.5		2325	10	US-09-775-879-22	Sequence 22, Appl
4	32.6	18.1		1830121	9	US-10-339-960-1	Sequence 1, Appl
5	32.4	18.3		353	10	US-09-560-863-384	Sequence 384, Appl
6	30	16.7		11316	9	US-09-764-868-1391	Sequence 1391, Appl
7	29.6	16.4		919	10	US-09-764-877-3680	Sequence 3680, Appl
8	29.6	16.4		3810	10	US-09-880-107-3709	Sequence 3709, Appl
9	28.8	16.3		1691139	9	US-10-067-514-1	Sequence 1, Appl
10	28.6	15.3		1260	9	US-09-938-842A-1758	Sequence 1758, Appl
11	28.4	15.3		1638	10	US-09-879-036-1	Sequence 1, Appl
12	27.8	15.4		277	10	US-09-923-876-5279	Sequence 5279, Appl
13	27.8	15.4		1977	9	US-09-738-626-3315	Sequence 3315, Appl
14	27.8	15.4		3309400	9	US-09-738-626-1	Sequence 1, Appl
15	27.6	15.3		682	9	US-10-206-901B-48	Sequence 48, Appl
16	27.6	15.3		2276	10	US-09-927-738-27	Sequence 27, Appl
17	27.2	15.1		474	9	US-09-918-995-1284	Sequence 1284, Appl
18	27.2	15.1		478	9	US-09-918-995-24377	Sequence 24377, Appl
19	27.2	15.1		1761	9	US-09-938-842A-303	Sequence 303, Appl

RESULT 2

US-09-775-879-20
; Sequence 20, Application US/09775879
; Patent No. US20020068822A1
; GENERAL INFORMATION:
; APPLICANT: Choe, Sunghwa
; APPLICANT: Feldmann, Kenneth A
; TITLE OF INVENTION: Dwf7 MUTANTS
; FILE REFERENCE: 2225-0003
; CURRENT APPLICATION NUMBER: US/09/775,879
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/179,901
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1889
; TYPE: DNA
; ORGANISM: Genomic dwf7 (Arabidopsis)
US-09-775-879-20

Query Match 100.0%; Score 180; DB 10; Length 1889;
Best Local Similarity 100.0%; Pred. No. 1.8e-53;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCGCGGATAATGCTTATCTGATGCAGTTTGTGACGAAACCTCTTTTACACCGA	60
DB	143	ATGGCGCGGATAATGCTTATCTGATGCAGTTTGTGACGAAACCTCTTTTACACCGA	202
QY	61	ATCGTTCTGAGTCATCTTTTGGCGGGAATCATGCGGAACCTTACCTCATTTCTCCAG	120
DB	203	ATCGTTCTGAGTCATCTTTTGGCGGGAATCATGCGGAACCTTACCTCATTTCTCCAG	262
QY	121	ACATGCTCCGAAATACCTTCGCGGAACCTTACTATCTTCTATCTCCGGTTCTCTCG	180
DB	263	ACATGCTCCGAAATACCTTCGCGGAACCTTACTATCTTCTATCTCCGGTTCTCTCG	322

RESULT 3

US-09-775-879-22
; Sequence 22, Application US/09775879
; Patent No. US20020068822A1
; GENERAL INFORMATION:
; APPLICANT: Choe, Sunghwa
; APPLICANT: Feldmann, Kenneth A
; TITLE OF INVENTION: Dwf7 MUTANTS
; FILE REFERENCE: 2225-0003
; CURRENT APPLICATION NUMBER: US/09/775,879
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/179,901
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 2925
; TYPE: DNA
; ORGANISM: Genomic HDF7
US-09-775-879-22

Query Match 69.6%; Score 125.2; DB 10; Length 2925;
Best Local Similarity 81.5%; Pred. No. 5.6e-34;
Matches 145; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY	3	GCGCGGATAATGCTTATCTGATGCAGTTTGTGACGAAACCTCTTTTACACCGAAT	62
DB	1511	GCGGATATGGAGATTAATGATCATGATCGTAATGAGACCTCTTTTACACCGAAT	1570
QY	63	CGTTCAGTCATCTTTTCCGCGGAATCATGCGGAACCTTACCTCATTTCTCCAGAC	122
DB	1571	GGTTCGAGTCACCTTTTCCGCGGAATCATGCGGAACCTTACCTCATTTCTCCAGAC	1630

QY	123	ATGGCTCCGAAATACCTCCGCGGAACCTACTATATCTATCTATCTATCTCGGTTTCTCTGG	180
DB	1631	ATGGCTCCGAACTACCTCCGCGGAACATACTCTACTTCACTCCGCTTCTCTGG	1688

RESULT 4

US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, F
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```

; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107246)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature

```

```

1 LOCATION: (120038)..(120038)
2 OTHER INFORMATION: n equals a, t, g or c
3 FEATURE:
4 NAME/KEY: misc_feature
5 LOCATION: (121344)..(121344)
6 OTHER INFORMATION: n equals a, t, g or c
7 FEATURE:
8 NAME/KEY: misc_feature
9 LOCATION: (122167)..(122167)
10 OTHER INFORMATION: n equals a, t, g or c
11 FEATURE:
12 NAME/KEY: misc_feature
13 LOCATION: (122336)..(122336)
14 OTHER INFORMATION: n equals a, t, g or c
15 FEATURE:
16 NAME/KEY: misc_feature
17 LOCATION: (131340)..(131340)
18 OTHER INFORMATION: n equals a, t, g or c
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (131360)..(131360)
22 OTHER INFORMATION: n equals a, t, g or c
23 FEATURE:
24 NAME/KEY: misc_feature
25 LOCATION: (139910)..(139910)
26 OTHER INFORMATION: n equals a, t, g or c
27 FEATURE:
28 NAME/KEY: misc_feature
29 LOCATION: (140398)..(140398)
30 OTHER INFORMATION: n equals a, t, g or c
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: (142750)..(142750)
34 OTHER INFORMATION: n equals a, t, g or c
35 FEATURE:
36 NAME/KEY: misc_feature
37 LOCATION: (145058)..(145058)
38 OTHER INFORMATION: n equals a, t, g or c
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: (145171)..(145171)
42 OTHER INFORMATION: n equals a, t, g or c
43 FEATURE:
44 NAME/KEY: misc_feature
45 LOCATION: (145942)..(145942)
46 OTHER INFORMATION: n equals a, t, g or c
47 FEATURE:
48 NAME/KEY: misc_feature
49 LOCATION: (147197)..(147197)
50 OTHER INFORMATION: n equals a, t, g or c
51 FEATURE:
52 NAME/KEY: misc_feature
53 LOCATION: (150841)..(150841)
54 OTHER INFORMATION: n equals a, t, g or c
55 FEATURE:
56 NAME/KEY: misc_feature
57 LOCATION: (152500)..(152500)
58 OTHER INFORMATION: n equals a, t, g or c
59 FEATURE:
60 NAME/KEY: misc_feature
61 LOCATION: (152530)..(152530)

```

Query Match 18.1%; Score 32.6; DB 9; Length 1830121;
Best Local Similarity 54.6%; Pred. No. 6.6;
Matches 65; Conservative 0; Mismatches 54; Indels 0;

Qy 59 GAATCGTTCTGAGTCATCTTTTGGCGGGGAATCTATGGGAACCCTTACCTCATTTTCTCC 118

Db 1431803 GAGCCATACAGAGACATCTTTGGGTGCAGAAAATACCCGCACTCATAATTTATTACCCA 1431862

0v
119 ACACATCCCTCCGAGAAATTACCCTCCCCCAGCCCTACTATACATCAATCATCCCCTTTCCTTC 177

[illegible]

```

RESULT 5
US-09-560-863-384/c
; Sequence 384, Application US/09560863
; Patent No. US20020110809A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael C.
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020110809A1 Human Polynucleotides and the
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: LEX-0018-USA
; CURRENT APPLICATION NUMBER: US/09/560,863
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/132,408
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 384
; LENGTH: 353
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-560-863-384

```

Qy	7	GCGGATTAATGCTTATCTGATGCAGTTGTGTGACGAACCCCTCTTTTTACAAACCGAATCGTT	66
Dd	215	GCGAAGAATCTTTTCGTGCTGTTGTGTTTCCAAAACCCTTCGAAATGCATTATCTTC	156
Qy	67	CTGAGTCATCTTTT	80
Dd	155	TTGTTGTTATCAATT	142

RESULT 7
US-09-764-877-3680
; Sequence 3680, Application US/09764877

```

; Patent No. US20020147140A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Rosen et al.
;
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;
; FILE REFERENCE: PC005
;
; CURRENT APPLICATION NUMBER: US/09/764,877
;
; CURRENT FILING DATE: 2001-01-17
;
; Prior application data removed - refer to PALM or file wrapper
;
; NUMBER OF SEQ ID NOS: 4031
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 3680
;
; LENGTH: 919
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-09-764-877-3680

```

```

RESULT 8
US-09-880-107-3709
: Sequence 3709, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880.107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3709
: LENGTH: 3810
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 X16349
: US-09-880-107-3709

```

[illegible]

	Query Match	15.9%	Score 28.6	DB 9	Length 1260
	Best Local Similarity	61.3%	Pred. No. 6.4		
	Matches 46	Conservative 0	Mismatches 29	Indels 0	Gaps 0
QY	104	TACTTCATTTTCCAGACATGCTCCGAATTAACCTGCCGGAAGCCTACTACTTCA	163		
Db	926	TTCCCTCAGTATCTCCCTCTCTATGTGCCACAAATTTCTCCTTGTAACCCCTGCTTGAGTTCA	867		
QY	164	TCGCCGGTTTCCCTCT	178		
Db	866	TGTTTCAGTCTCTGT	852		

```

RESULT 12
US-09-923-876-5279
; Sequence 5279, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND P
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5279
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incvte ID No. US20020

```


Result No.	Score	Query Match	Length	DB	ID	Description	
1	71.2	39.6	459	4	US-09-443-041A-13	Sequence 13, Appl	
2	71.2	1447	1447	4	US-09-443-041A-27	Sequence 27, Appl	
3	67.6	37.6	1318	4	US-09-443-041A-25	Sequence 25, Appl	
4	58	32.3	1070	4	US-09-443-041A-9	Sequence 9, Appl	
5	56.2	31.3	360	4	US-09-443-041A-11	Sequence 11, Appl	
6	37.4	20.8	425	4	US-09-443-041A-15	Sequence 15, Appl	
7	37.4	20.8	1210	4	US-09-443-041A-29	Sequence 29, Appl	
C 8	28.8	16.0	2022	1	US-08-803-973-6	Sequence 6, Appl	
C 9	28.8	16.0	2022	1	US-08-803-972-6	Sequence 6, Appl	
C 10	28.8	16.0	2124	1	US-08-803-973-11	Sequence 11, Appl	
C 11	28.8	16.0	2124	1	US-08-803-972-11	Sequence 11, Appl	
C 12	28.8	16.0	4371	1	US-08-803-973-1	Sequence 1, Appl	
C 13	28.8	15.0	4371	1	US-08-803-972-1	Sequence 1, Appl	
C 14	28.2	15.7	87350	3	US-08-781-891-79	Sequence 79, Appl	
C 15	28.2	15.7	87543	4	US-09-791-211-3	Sequence 3, Appl	
C 16	28	15.6	7218	1	US-08-232-463-14	Sequence 14, Appl	
C 17	28	15.6	35100	2	US-08-770-379-19	Sequence 19, Appl	
C 18	28	15.6	35100	4	US-08-757-669A-19	Sequence 19, Appl	
C 19	28	15.6	35100	4	US-09-230-371A-19	Sequence 19, Appl	
C 20	27	15.6	575	4	US-09-328-111-325	Sequence 325, App	
C 21	27	15.6	726	4	US-08-136-164-72	Sequence 72, Appl	
C 22	27	15.6	8056	3	US-09-136-605-14	Sequence 14, Appl	
C 23	26.8	14.5	1387	4	US-09-461-474-1	Sequence 1, Appl	
C 24	25.6	14.2	4047	2	US-08-612-734B-1	Sequence 1, Appl	
C 25	25.6	14.2	4800	2	US-08-612-734B-3	Sequence 3, Appl	
C 26	25.6	14.2	176373	3	US-09-128-155-17	Sequence 17, Appl	
C 27	25.4	14.1	1590	4	US-09-134-001C-1460	Sequence 1460, App	


```
; ORGANISM: Oryza sativa
US-09-443-041A-11

Query Match      31.2%; Score 56.2; DB 4; Length 360;
Best Local Similarity 59.9%; Pred. No. 1.2e-10;
Matches 94; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 19 TATCTATGCGAATTTGTTGACGAAACCTCTTTTACAAACCGAATCGTTCTGAGTCATCTT 78
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 TACCTTCGCCAGTTTCGTCGAGGAGAGCGCTGGTACACGAGATCTTCTCAGCCATGTG 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 79 TTCCGTCGGAATCTATGGAACCTTACTCTATTTTCCAGACATGCTCGGAATATAC 138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 GTCCGTCGCGACTGGTGGCGCGCTCCCTCCACCGCTCCAGTCGTGGCTCGCAACGGT 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 139 CTCGCGTGAACCTCTACTTACTTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 CTCGCGTGAACCTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-443-041A-15
; Sequence 15, Application US/09443041A
; Patent No. 646571
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (286)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (288)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (328)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (344)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (372)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (390)..(391)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (450)..(451)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (480)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (489)
; OTHER INFORMATION: any nucleotide
US-09-443-041A-15

Query Match      20.8%; Score 37.4; DB 4; Length 495;
Best Local Similarity 64.4%; Pred. No. 0.00059;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 94 TGGGAACCTTACCTCATTTTCTCCAGACATGCTCCGAAATTAATCTCCCGGAACCTTA 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 50 TGGCGGGCGCTGCCGACCGCTCCGCTCTCTGGCTGCGCAACTGCATCGCGGCTACCTC 109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 154 CTATACCTTCACTCTCCGGTTTCTCTCTGG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 CTCTACTTGGCCACCGGCTTCTCTCTGG 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-443-041A-29
; Sequence 29, Application US/09443041A
; Patent No. 646571
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1129
; OTHER INFORMATION: any nucleotide
US-09-443-041A-29

Query Match      20.8%; Score 37.4; DB 4; Length 1210;
Best Local Similarity 64.4%; Pred. No. 0.00085;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 94 TGGGAACCTTACCTCATTTTCTCCAGACATGCTCCGAAATTAATCTCCCGGAACCTTA 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 TGGCGGGCGCTGCCGACCGCTCCGCTCTCTGGCTGCGCAACTGCATCGCGGCTACCTC 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 154 CTATACCTTCACTCTCCGGTTTCTCTCTGG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 CTCTACTTGGCCACCGGCTTCTCTCTGG 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-803-973-6/c
; Sequence 6, Application US/08803973
; Patent No. 5773701
; GENERAL INFORMATION:
; APPLICANT: Braun, Carl J.
; APPLICANT: Hemenway, Cynthia L.
; APPLICANT: Tumer, Nilgun E.
; TITLE OF INVENTION: Plants Resistant to Infection by PVX
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

; INFORMATION FOR SEQ ID NO: 11:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 2124 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;         NAME/KEY: CD3
;         LOCATION: 1..2124
US-08-803-972-11

```

Query Match	16.0%;	Score 28.8;	DB 1;	Length 4371;
Best Local Similarity	50.0%;	Pred. No. 1.6;		
Matches	48;	Conservative 0;	Mismatches 32;	Indels 0; Gaps 0;
QY	27	GCATTTGTTGACGAACCTCTTTTTACAACCGAATCGTTCAGCACTACTTTTGCCGGC	86	
Db	1883	GCATTTAGTTGAGGAATCCAGTGTTCACAGGAGCTCTCTATGACTGTTGCCGTCGT	1824	
QY	87	GAATCTATGGGAACCCCTTAC	106	

Db 1823 GTCTCAACGGCTCTTTTC 1804

RESULT 13

US-08-803-972-1/c

Sequence 1, Application US/08803972

Patent No. 5792937

GENERAL INFORMATION:

APPLICANT: Braun, Carl J.

APPLICANT: Hemenway, Cynthia L.

APPLICANT: Turner, Nilgun E.

TITLE OF INVENTION: Plants Resistant to Infection by PVX

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B4F

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/803,972

FILING DATE: 21-FEB-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/804,862

FILING DATE: 04-DEC-1991

APPLICATION NUMBER: US 07/771,912

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10541)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6099

TELEFAX: (314)537-6047

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4371 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..4371

US-08-803-972-1

Query Match 16.0%; Score 28.8; DB 1; Length 4371;

Best Local Similarity 60.0%; Pred. No. 1.6;

Matches 48; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 GCAGTTGTTGACGAACCTCTTTTACAAACCGAATCGTTCTGAGTCATCTTTTGCCGCGC 86

Db 1883 GCATTTAGTAGGAAATCCAGTGTTCACAGGAGTCCTCTATGACTGTGTCGCGCT 1824

QY 87 GAATCTATGGGAACCTTAC 106

Db 1823 GTCTCAACGGCTCTTTTC 1804

RESULT 14

US-08-781-891-79

Sequence 79, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,891

FILING DATE: 27-DEC-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: NO. 6090620tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 87350 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-79

Query Match 15.7%; Score 28.2; DB 3; Length 87350;

Best Local Similarity 54.3%; Pred. No. 8.8;

Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 11 ATAATGCTTATCTGATGCGAGTTGTGTGACGAACCTCTTTTACACCGAATCGTTCTGA 70

Db 69319 ATAATACGTTTATGTGTACCTTGTATTATTATTATGATAGTAATAATTATAC 69378

QY 71 GTCATCTTTTGGCGGAATCTATGGAACCTTACCTCATTTTC 115

Db 69379 AACACAGATATGCTCTAATCACTGGGAGCCCTGAGCTTGTTC 69423

RESULT 15

US-09-791-211-3

Sequence 3, Application US/09791211

Patent No. 6448080

GENERAL INFORMATION:

APPLICANT: Donna T. Ward

APPLICANT: Andrew T. Watt

TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION

FILE REFERENCE: RTS-0205

CURRENT APPLICATION NUMBER: US/09/791,211

CURRENT FILING DATE: 2001-02-23

NUMBER OF SEQ ID NOS: 90

SEQ ID NO 3

LENGTH: 87543

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: 7421

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 7427

NAME/KEY:	unsure	unknown
LOCATION:	47291	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	52786	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	52787	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	53384	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	54684	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	59215	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	59235	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	59242	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	63290	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	66614	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	68660	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	68697	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	68718	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	68733	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	68739	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	69785	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	79134	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	79198	unknown
OTHER INFORMATION:		unknown
NAME/KEY:	unsure	unknown
LOCATION:	86336	unknown
OTHER INFORMATION:		unknown
OTHER INFORMATION:		unknown
US-09-791-211-3		

Query Match	15.7%	Score 28.2;	DB 4;	Length 87543;
Best Local Similarity	54.3%	Pred. No. 8.8;		
Matches	57;	Conservative	0;	Mismatches 48; Indels 0;
Qy	11	ATAATGCTTATCTGATCGAGTTTCTTGTGACGAAACCTCTTTTACACCGCAATC		
Db	69512	ATAATACGTTTATGTCTGCTCTTATTATTATACATTGTAGTATAGAATA		
Qy	71	GTCATCTTTTGGCCGGCGAATCTATGGGAACCTTACCTCATTTTC		115
Db	69572	AACACAGCAATATGCTCAATCAGTGGGAGCCCTGAGTTGTTTTC		69616

Search completed: June 6, 2003, 16:28:03
Job time : 37.9245 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:20 ; Search time 423 Seconds
(without alignments)
7506.661 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552

Perfect score: 1410

Sequence: 1 at:gcgcgcgataatgctta.....tgcatcatcgcaacatctg 1410

Scoring table: IDENTITY_NUC
Gap: 10.0 , Gapext 1.0

Searched: 2185339 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278.4	19.7	1164	21 AAC41897	Arabidopsis thaliana
2	164.2	11.5	1381	22 AAH50964	Sterol C5 desaturase
3	158.2	11.3	446	21 AAG69689	Eucalyptus grandis
4	125.4	8.3	285	24 ABL72194	Corn tassell-derivative
5	106.2	7.5	304	24 ABL70846	Corn tassell-derivative
6	54.8	3.3	6485	22 AAS46559	Tumour suppressor
7	54.8	3.3	6485	24 ABL33808	Human immune system
8	54.2	3.3	6015	24 ABL32676	Human immune system
9	54	3.3	7008	22 AAS46531	Tumour suppressor

c	10	53.6	3.8	875	22	AAI95044	Human neuroblastoma
	11	52	3.7	9539	22	AAS45347	Chemically pretreated DNA transcription
	12	52	3.7	9539	24	ABK28180	Human immune system
	13	51.4	3.6	6145	24	ABL32972	Chemically treated
	14	51.2	3.6	5181	24	ABL70443	Human gene regulated
	15	51.2	3.6	5493	24	AAS61386	Human immune system
	16	51.2	3.6	12007	24	ABL32717	Oligonucleotide of
	17	51	3.6	693	24	ABQ32104	Oligonucleotide of
c	18	51	3.6	6301	24	ABQ32053	Human immune system
	19	50.8	3.6	6301	24	ABQ32053	Human immune system
	20	50.4	3.6	556	24	ABQ36996	Oligonucleotide of
c	21	50.4	3.6	556	24	ABQ36997	Oligonucleotide of
	22	50	3.5	5432	24	ABN80000	Human chemically modified
c	23	49.8	3.5	6644	20	AAX33181	Base sequence of tumour suppressor
	24	49.8	3.5	7118	22	AAS46381	Tumour suppressor
	25	49.8	3.5	7118	24	ABL32830	Human immune system
	26	49.8	3.5	7348	22	AAS46335	Tumour suppressor
c	27	49.8	3.5	7372	20	AAX33182	Base sequence of tumour suppressor
	28	49.8	3.5	7797	20	AAX33180	Cowpox virus base sequence
c	29	49.8	3.5	7996	20	AAX33184	Base sequence of tumour suppressor
	30	49.4	3.5	5269	24	ABL34056	Human immune system
	31	49.4	3.5	12142	24	ABL33673	Human immune system
	32	49.4	3.5	14006	24	ABL33958	Human immune system
	33	49	3.5	47108	24	ABK31510	Signal transduction
	34	48.8	3.5	5586	24	ABK40004	Human chemically pretreated
	35	48.8	3.5	6334	24	ABL33213	Human immune system
	36	48.8	3.5	6668	24	ABL33697	Human immune system
	37	48.8	3.5	17594	24	ABL34026	Human immune system
	38	48.8	3.5	17934	24	ABL33719	Human immune system
	39	48.4	3.4	5437	24	ABL33412	Human immune system
	40	48.4	3.4	12465	22	AAS45326	Chemically pretreated
	41	48.4	3.4	12465	24	ABL32634	Human immune system
c	42	48.2	3.4	819	22	AAI94974	Human neuroblastoma
	43	48.2	3.4	6656	24	ABL33400	Human immune system
c	44	48	3.4	626	23	ABV60941	Human prostate expression
	45	48	3.4	6794	24	ABL70220	Chemically treated

ALIGNMENTS

RESULT 1

AAC41897
ID AAC41897 standard; DNA; 1164 BP.
XX AAC41897;
AC AAC41897;
XX AAC41897;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 33540.
XX Arabidopsis thaliana.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PT cancer -
PS Claim 1; SEQ ID NO 281; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC fcp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6485 BP; 1509 A; 168 C; 1822 G; 2985 T; 1 other;

Query Match 3.9%; Score 54.8; DB 22; Length 6485;
Best Local Similarity 46.2%; Pred. No. 0.00068;
Matches 182; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 191 TCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCGACCTTCACCTTTGTATTC 250
Db 2363 TTTTATAGTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2422
QY 251 ACTATTGCTTAATCGCTTCTATGTATCGATTTTCAATTTAAGGAAGAGGTTTCTTC 310
Db 2423 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTT 2482
QY 311 GTGCTACTGACAGTAATTTGGATTTCATGTGGATTCGATTCGATTTTCAATTTAAGGAAGAGGTTTCTTC 370
Db 2483 ATTTTGTGTTT 2542
QY 371 TTGTGCATATCTCCATCTAAGGGATTGAACAGTTAGTGGCTTATATAAGTTTCTGCA 430
Db 2543 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2602
QY 431 ACCAATGAGAAGTCGTACATCTTTGAAGTTTCAATTTTCTACTTGCATTTAAGTCCACTT 490
Db 2603 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2662
QY 491 AAATGTTTGTGGAAGTGAATGCTACTTTCAGACACATCTTTTCTGCTCTCGAGA 550
Db 2663 TTAATT 2722
QY 551 CTCGTGCTTAGTTTGAATCTTTTGGTCTGTT 584
Db 2723 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2756

RESULT 7

ABL33808

ID ABL33808 standard; DNA; 6485 BP.

XX

AC ABL33808;

XX

DT 26-MAR-2002 (first entry)

XX

DE Human immune system associated gene SEQ ID NO: 1781.

XX

KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosine; cytosine; cytosine;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW antineuritic; antiarthritic; antidiabetic; antipsoriatic;
KW antineuritic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW ds.
XX Homo sapiens.
XX WO200200928-A2.
PN 03-JAN-2002.
XX 02-JUL-2001; 2001WO-EP07537.
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX Claim 1; SEQ ID NO 1781; 32pp + Sequence Listing; German.
PS The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX Sequence 6485 BP; 1509 A; 168 C; 1822 G; 2985 T; 1 other;

Query Match 3.9%; Score 54.8; DB 24; Length 6485;
Best Local Similarity 46.2%; Pred. No. 0.00068;
Matches 182; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 191 TCTATTACCTTAAATCAACGTTTACCTTCCCAAGGCTCGACCTTCACCTTTGTATTC 250
Db 2363 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2422
QY 251 ACTATTGCTTAATCGCTTCTATGTATCGATTTTCAATTTAAGGAAGAGGTTTCTTC 310
Db 2423 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2482
QY 311 GTGCTACTGACAGTAATTTGGATTTCATGTGGATTCGATTCGATTTTCAATTTAAGGAAGAGGTTTCTTC 370
Db 2483 ATTTTGTGTTT 2542
QY 371 TTGTGCATATCTCCATCTAAGGGATTGAACAGTTAGTGGCTTATATAAGTTTCTGCA 430
Db 2543 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2602
QY 431 ACCAATGAGAAGTCGTACATCTTTGAAGTTTCAATTTTCTACTTGCATTTAAGTCCACTT 490
Db 2603 TTTTATAGTTTTTAAAGTTTTTATTCGTTTGTGATATGTTTTTCGTTTTTTTTTTT 2662
QY 491 AAATGTTTGTGGAAGTGAATGCTACTTTCAGACACATCTTTTCTGCTCTCGAGA 550
Db 2663 TTAATT 2722
QY 551 CTCGTGCTTAGTTTGAATCTTTTGGTCTGTT 584


```
Db 2723 TTTTTFMTTTTTTTTTTTTATTATTATTCGTGTT 2756
RESULT 8
ABL32676
ID ABL32676 standard; DNA; 6015 BP.
XX
AC ABL32676;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 649.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX
PS Claim 1; SEQ ID NO 649; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6015 EP; 1822 A; 96 C; 1277 G; 2820 T; 0 other;
Query Match 3.8%; Score 54.2; DB 24; Length 6015;
Best Local Similarity 46.6%; Pred No. 0.00095;
Matches 173; Conservative 0; Mismatches 198; Indels 0; Gaps 0;
QY 243 TTGTAITACATTCGCTTAATCGCTTCTATGTTATCGATTTTCAATTTAAGGAGAGG 302
Db 1012 TTGTCGTTTTTTTTTTTTTTTATTTTATTTTGTATATATTTTATTTGATATATTTG 1071
QY 303 GTTCTTCGTACTGACAGTAATTTGGATTTGATGGAGTAGTTCATGTTTGCATTTA 362
Db 1072 TGGGGTATGTGAAATTTGTGTATGTATAGAAATGCTATGATTAAGTTAGTAGATCTT 1131
QY 363 TTGATATTGTCATATTCCTCACTAAGGAGTTGAACAGTTAGTGCCTTATATACTT 422
Db 1132 GGTATPTTTATTTATTTATTTTATTTATGTTGTGATATATTTTAAAGTTTTTTTGT 1191
QY 423 TTGTGCAACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTTCTACTTGCCATTTAA 482
Db 1192 ATTTAGAAAATTATAATATTTTGTGTTATTAGTTAGTTATTTATTTTGTTAATAT 1251
QY 483 GTCCACTTAAATTTGTTTGAAGTGAATGCTACTTTTCACACACATTCCTTTCTGCTT 542
Db 1252 TAGAAAGTATTTTATTTTAAATTTTGTGTTGTTATTTTATTAATATTTTTCGTT 1311
QY 543 CTCTGAGACTCTGCTTGAATCTTTTGGTGTGTTTCTTCTTTCAGATGCAATTC 602
Db 1312 TTTATTTTATTTTATATATATATATTTTATTTAGTTTTCGTAATATATATTTT 1371
QY 603 CTACAATAAAG 613
Db 1372 TTACGGTTATG 1382
RESULT 9
AAS46531
ID AAS46531 standard; DNA; 7008 BP.
XX
AC AAS46531;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #253.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
XX
PR 06-APR-2000; 2000DE-1019058.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
PS Claim 1; SEQ ID No 253; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and
XX oncogenes having a sequence taken from 536 (actually 533 since
XX numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX (SS) and sequences complementary to (SS). The nucleic acid may be a
XX peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX form part of a set of probes for detecting the cytosine methylation state
XX and/or single nucleotide polymorphisms and also to be used in an
XX array for analysing diseases associated with CpG dinucleotides e.g.
XX cancers and tumours. The probes can also be used in a method for
XX ascertaining genetic and/or epigenetic parameters for the diagnosis
XX and/or therapy of existing diseases or the predisposition to specific
XX diseases, by analysing cytosine methylations. The parameters may be
XX compared to another set of genetic and/or epigenetic parameters, the
XX differences serving as basis for diagnosis and/or prognosis events which
```

are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

SQ Sequence 7008 BP; 2322 A; 26 C; 1306 G; 3354 T; 0 other;
Query Match 3.8%; Score 54; DB 22; Length 7008;
Best Local Similarity 46.0%; Pred. No. 0.0011;
Matches 183; Conservative 0; Mismatches 215; Indels 0; Caps 0;

Qy	202	AAATCAACGCTTTACCTTCCCAAGGCTCGACCTTCACATTTGTGATTCACATATTCGTTA	261
Db	3316	AAATTTAAAGTTTTATTTATTTAGATTTGATTTAGATTTTATTTGTTTTATTTATTTT	3375
Qy	262	ATCGCTTTCATGTTATPCGATTTTTCAAATTTAAGGAAGAGGGTTCTTCGCGACATCTAC	321
Db	3376	ATTTTATTTTGTGTTTTTATTTATTTGTTTAAAGTTAGAGTGTTTTTTTGGATGTTTTTT	3435
Qy	322	AGTAAATTTGGATTTGATGTCGATAGTTTCATGTTTGGCATTTATTGATTTATTTGCCATATT	381
Db	3436	AAATATGTTAAATTAGCGTTAGTTTAAAGATTTTTCGGTTTATTAATTTTTTTTGTGTTGA	3495
Qy	382	CTCCATCTAAGGATTTGAACAGTTAGTGGCTTATATAAAGTTTTTTGTGCAACCAATGAGAA	441
Db	3496	ATGTATATTTTTTGGAAATAAATGTATAGTTTTAAATTTTTTATTTTTTTTATGTTTTGTGTG	3555
Qy	442	GTGCTACATCTTTGAAGTTGAAATTTCTACTTGCCATTTAAGTCCACTTAAATGTTTGT	501
Db	3556	ATATATTTGAAAAGTAAGATTTTTTTTGAATACGTTATGATATATATTTTTTTTAATTTT	3615
Qy	502	TGAAGTGATTTGCTACTTTTCAGACACATCTTTTTCTGCTTCTCTGAGACTCTGCTTAG	561
Db	3616	TTAGTATTTTATTTTTTTTGTGTTTTTTTTTATATAATTTATTTTGNATATATTTTATAAT	3675
Qy	562	TTTGAATCTTTTTTGGTCTGTTTTGCTTCAGATGCAA	599
Db	3676	TATTTATTTATTTTTTGGTATTTCTGTTTTTTTTTTATATAAAA	3713

RESULT 10	
AAI95044/c	
ID	AAI95044 standard; cDNA; 875 BP.
XX	
XX	
AC	AAI95044;
XX	
DT	13-NOV-2001 (first entry)
XX	
DE	Human neuroblastoma expressed polynucleotide SEQ ID NO 1119.
XX	
KW	Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200166719-A1.
XX	
XX	
PD	13-SEP-2001.
XX	
XX	
PF	02-MAR-2001; 2001WO-JP01629.
XX	
PR	07-MAR-2000; 2000JP-0159195.
XX	
XX	
PA	(CHIB-) CHIBA PREFECTURE.
PA	(HISM) HISAMITSU PHARM CO LTD.
XX	
XX	
PI	Nakagawara A;
XX	
DR	WFI; 2001-565584/63.
XX	
XX	
PT	Nucleic acids originating in gene expressed in human neuroblastoma,

PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX
XX Claim 1; Page 853; 2979pp; Japanese.
PS
XX
XX The invention relates to novel genes (AAI93126-AAI97963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX
XX Sequence 875 BP; 619 A; 73 C; 40 G; 25 T; 118 other;
SQ

Query Match	3.8%;	Score 53.6;	DB 22;	Length 875;
Best Local Similarity	41.5%;	Pred. No. 0.00069;		
Matches	146;	Conservative 0;	Mismatches 206;	Indels 0; Gaps 0;
Qy	235	TTTCACATTTTGTATTCACACTATTGCTTAATACGCTTCTCATGTTATCGATGATTTTCAATTAA	294	
Db	545	TTTTNTNTTTTTTGTGTTTNCNTNTTTTTNTNTTTTTNTTTTTNTTTTNCNTNTTTTTNT	486	
Qy	295	GGAAGAGGGTTCCTTCGTACTGTCACAGATAAATTTGGATTTGATGTGGATAGTTCATGTT	354	
Db	485	NNTNCNNTTTTCTTTTNTGGTTTTTTTTTATNTNTTTTTTTTTNTGTGTTTTNTTTT	426	
Qy	355	TGCATTTATGATTAATTTGTGCATATTTCTCCATCTAAGGATGAACAGTTAGTGGCTTA	414	
Db	425	GTITTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGNNNTTTTTTTTTTTTTTTATNTTT	366	
Qy	415	TATAAGTTTTTGTGCAACCAATGAGAAGTCGTACATCTTTGAGTTTGAATTTCTACTTG	474	
Db	365	TTTTTTTTTTTTTTTTTTTGTGTTTGTGTTTTTTTTTTTTTTTTTTTGTGTTTATTTTGTGTTT	306	
Qy	475	CCATTTAAGTCCACTTAAATTTGTTGTTGAAGTGATTTGCTACTTTTCAGACACATTTCTTT	534	
Db	305	TTNTGTTNTTTTTTTTTTTTTTTTTTTTTTTTNNCTTTTTTTTTTGGTTTTTTTTTNCNTGTTTTTTT	246	
Qy	535	TTTCGCTCTCTGAGACTCTGCTCTACTGTTGGAATCTTTTTTGGTCTGTTTTT	586	
Db	245	TTTTNTTTTTGTGNTTCTNTNTNTNGTTTTTTTTTTNGTTTTTTTTTGTGTTTGTGTTT	194	

RESULT	11	
AAAS45347		
ID	AAAS45347	standard; DNA; 9539 BP.
XX	XX	
AC	AAAS45347;	
XX	XX	
DT	18-DEC-2001	(first entry)
XX	XX	
DE	XX	Chemically pretreated complementary DNA associated with cell cycle #26.
XX	XX	Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
KW	KW	human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW	KW	graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
KW	KW	arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
KW	KW	immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
XX	XX	PCR primer.
XX	XX	
OS	OS	Homo sapiens.
XX	XX	
FN	WO200168911-A2.	
XX	XX	
PD	20-SEP-2001.	
XX	XX	
PF	15-MAR-2001; 2001WO-EP02945.	
XX	XX	
PR	15-MAR-2000; 2000DE-1013847.	
PR	06-APR-2000; 2000DE-1019058.	
PR	07-APR-2000; 2000DE-1019173.	
PR	30-JUN-2000; 2000DE-1032529.	

PR 01-SEP-2000; 20(ODE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602751/68.
XX Designing primers and probes for analysing diseases associated with
PT cytosine methylation state e.g. arthritis, cancer, aging,
PT arteriosclerosis comprising fragments of chemically modified genes
PT associated with cell cycle
XX
PS Claim 1; SEQ ID No 52; 28pp; English.
XX
XX Sequences AAS45:96-AAS45520 represent chemically pretreated genomic DNA
CC molecules associated with the cell cycle and specific PCR primers of the
CC invention. The sequences are useful for detecting the methylation state
CC of all CpG dinucleotides in a sequence and therefore for analysing
CC associated diseases. By analysing cytosine methylations in the pretreated
CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
CC of existing diseases or the predisposition to specific diseases can be
CC ascertained. The parameters may be compared to another set of genetic
CC and/or epigenetic parameters, the differences serving as basis for
CC diagnosis and/or prognostic events which are disadvantageous to patients.
CC The sequences of the invention are useful for the diagnosis and therapy
CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
CC aging, glomerular disease, Lewy body disease, arthritis,
CC arteriosclerosis, solid tumours and cancers.
XX
SQ Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 other;
Query Match 3.7%; Score 52; DB 22; Length 9539;
Best Local Similarity 46.6%; Pred. No. 0.0041;
Matches 166; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 235 TTTCAC TTGTTGTTACATGCTGTTAATCGCTTTCTATGTTATCGATTTCATTTAA 294
Db III III III III III III III III III III III III III III III
55 TTATTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTATTATTAGTTTTTATA 114
QY 295 GGAAGA JGGTTCTCGTACTGTTACAGTAATTTGGATTTCATGTCGATGTCATGTT 354
Db III III III III III III III III III III III III III III III
115 TTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 174
QY 355 TGCATTATGATTATTTGTCATATTTCTCCATCTAAGGGAATGAACAGTTAGTGGCTTA 414
Db III III III III III III III III III III III III III III III
175 TTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 234
QY 415 TATAAG TTTTGTGCAACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTCTACTTG 474
Db III III III III III III III III III III III III III III III
235 TGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 294
QY 475 CCATTAGTCCACTTAATGTTGTTGAAGTGAATGTCATCTTCAGACACATCTTT 534
Db III III III III III III III III III III III III III III III
295 TTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 354
QY 535 TTCTGCTCTCTGAGACTGCTGCTAGTTTGAATCTTTTGGTCTGTTTGGCTT 590
Db III III III III III III III III III III III III III III III
355 TTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 410
RESULT 12
ID ABK28180
XX ABK28180 standard; DNA; 9539 BP.
AC ABK28180;
XX
DT 23-APR-2002 (first entry)
XX DNA transcription associated complementary genomic DNA #27.
DE
XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW immunological disorder; Werner syndrome; developmental disorder;
KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KW angioneurotic edema; congenital heart disease; HDR syndrome; gene therapy;
KW polyglutamine disorder; solid tumour.
XX Unidentified.
XX WO200192565-A2.
XX
XX 06-DEC-2001.
XX
XX 06-APR-2001; 2001WO-EP03973.
XX
XX 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
WPI; 2002-090046/12.
PT New nucleic acids or oligomers, useful for diagnosing or treating
PT diseases associated with DNA transcription, e.g. immunological
PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid
PT tumours or cancer
XX Claim 1; SEQ ID No 54; 32pp; English.
XX The invention relates to a nucleic acid, which comprises a segment of the
CC chemically pretreated DNA of genes associated with DNA transcription from
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
CC to the chemically pretreated DNA of genes associated with DNA
CC transcription. The set of oligomer probes are useful for detecting the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status), e.g. adenosine deaminase
CC deficiency, viral infection, retroviral infection, Sezary syndrome,
CC haematological disorders, immunological disorders, Werner syndrome,
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
CC neurological disorders, neurodegenerative disorders, Waardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 other;
Query Match 3.7%; Score 52; DB 24; Length 9539;
Best Local Similarity 46.6%; Pred. No. 0.0041;
Matches 166; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 235 TTTCAC TTGTTGTTACATGCTGTTAATCGCTTTCTATGTTATCGATTTCATTTAA 294
Db III III III III III III III III III III III III III III III
55 TTATTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTATTATTAGTTTTTATA 114
QY 295 GGAAGA JGGTTCTCGTACTGTTACAGTAATTTGGATTTCATGTCGATGTCATGTT 354
Db III III III III III III III III III III III III III III III
115 TTTTTCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 174
QY 355 TGCATTATGATTATTTGTCATATTTCTCCATCTAAGGGAATGAACAGTTAGTGGCTTA 414

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 13:35:44 ; Search time 250.075 Seconds
(without alignments)
1729.135 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atgcgcgcgataatgctta.....tgcatccatgcgaacatctg 1410

Scoring table: IDENT:TY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44136; seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /3gn2_6/ptodata/1/1na/5A_COMB.seq.*
2: /3gn2_6/ptodata/1/1na/5B_COMB.seq.*
3: /3gn2_6/ptodata/1/1na/6A_COMB.seq.*
4: /3gn2_6/ptodata/1/1na/6B_COMB.seq.*
5: /3gn2_6/ptodata/1/1na/PCTUS_COMB.seq.*
6: /3gn2_6/ptodata/1/1na/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
result No.	Score	Query Match	Length	DB	ID	Description
1	184	13.0	1447	4	US-09-443-041A-27	Sequence 27, Appl
2	166	11.8	911	4	US-09-443-041A-23	Sequence 23, Appl
3	165.8	11.8	1070	4	US-09-443-041A-9	Sequence 9, Appl
4	163.4	11.6	1210	4	US-09-443-041A-29	Sequence 29, Appl
5	161.8	11.5	1318	4	US-09-443-041A-25	Sequence 25, Appl
6	92.8	6.6	495	4	US-09-443-041A-15	Sequence 15, Appl
7	89	6.3	459	4	US-09-443-041A-13	Sequence 13, Appl
8	65	4.6	360	4	US-09-443-041A-11	Sequence 11, Appl
9	55.6	3.5	7218	1	US-08-232-463-14	Sequence 14, Appl
C 10	42.8	3.0	5058	4	US-09-889-595-1	Sequence 1, Appl
C 11	39.6	2.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 12	38.2	2.7	43676	3	US-09-356-952-12	Sequence 12, Appl
C 13	37.2	2.6	2060	1	US-07-721-761A-31	Sequence 31, Appl
C 14	37.2	2.6	2060	1	US-07-978-687-31	Sequence 31, Appl
C 15	37.2	2.6	2060	5	PCT-US91-01750-2	Sequence 2, Appl
C 16	37.2	2.6	2060	5	PCT-US91-05801-31	Sequence 31, Appl
C 17	37.2	2.6	2081	5	PCT-US91-01750-3	Sequence 3, Appl
C 18	37.2	2.6	3440	1	US-08-471-791-27	Sequence 27, Appl
C 19	37.2	2.6	3440	5	PCT-US91-01746-27	Sequence 27, Appl
C 20	37	2.6	4817	1	US-07-951-715A-18	Sequence 18, Appl
C 21	37	2.6	4817	2	US-08-459-448A-18	Sequence 18, Appl
C 22	37	2.6	4817	3	US-08-459-595A-18	Sequence 18, Appl
C 23	37	2.6	4817	3	US-08-459-504B-18	Sequence 18, Appl
C 24	37	2.6	4817	3	US-08-459-444-18	Sequence 18, Appl
C 25	37	2.6	4817	4	US-09-547-422-18	GENERAL INFORMATION
C 26	36.4	2.6	658	4	US-08-998-416-595	GENERAL INFORMATION
C 27	36.2	2.6	731	4	US-08-887-534A-8	Sequence 8, Appl

28	36	2.6	573	4	US-08-936-165A-48	Sequence 48, Appl			
c 29	35.8	2.5	1622	4	US-08-949-246-3	Sequence 3, Appl			
c 30	35.6	2.5	2484	4	US-08-961-527-234	Sequence 234, App			
31	35.4	2.5	10968	2	US-08-680-327-2	Sequence 2, Appl			
32	35.4	2.5	10968	4	US-09-228-246-1	Sequence 1, Appl			
c 33	35.2	2.5	289	4	US-09-007-005-17	Sequence 17, Appl			
c 34	35.2	2.5	289	4	US-09-244-796-17	Sequence 17, Appl			
c 35	34.8	2.5	41708	4	US-09-470-512A-3	Sequence 3, Appl			
c 36	34.8	2.5	51259	3	US-08-781-891-209	Sequence 209, App			
c 37	34.6	2.5	327	4	US-09-889-595-7	Sequence 7, Appl			
c 38	34.6	2.5	3457	1	US-08-295-882-1	Sequence 1, Appl			
c 39	34.2	2.4	6243	2	US-09-056-075-1	Sequence 1, Appl			
c 40	34	2.4	51259	3	US-08-781-891-209	Sequence 209, App			
41	34	2.4	112132	4	US-09-741-150-3	Sequence 3, Appl			
c 42	33.8	2.4	2370	1	US-08-104-072B-7	Sequence 7, Appl			
c 43	33.8	2.4	2370	1	US-08-351-413-8	Sequence 8, Appl			
c 44	33.8	2.4	2370	2	US-09-025-583-8	Sequence 8, Appl			
c 45	33.8	2.4	9636	1	US-08-323-170B-1	Sequence 1, Appl			

ALIGNMENTS

RESULT 1
US-09-443-041A-27
; Sequence 27, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: Glycine max
US-09-443-041A-27

Query Match	13.0%	Score 184;	DB 4;	Length 1447;
Best Local Similarity	72.6%	Pred. No. 3.8e-42;		
Matches 238;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;
QY	592	AGATGCAATTCCTACAAATAAAGGCTATGCGTTTGGCAATGTTGTGGCAATGAAGGCTAT	651	
Db	316	AGATGCTATTCCCTCTCAAGAGCCATGCTCTTGCNAATATCTGTTGTATGAAGCCAT	375	
QY	652	GCCATGTPACACTCTTCTTCCAACTGTCTCCGAGAGTATGATTGAACCTGTTGGACCAA	711	
Db	376	GCCGTGGTACTCTTTACTTCCAACTGTTTCGGAGTACCTGGTAGAACTGGCTGGACAAA	435	
QY	712	ATGTTTCTAGCATAGACAGAAATTCGGTGGATCTGTATTTTGTATCATCGCCATCTA	771	
Db	436	GTGTATCTAGATATATATAATGTTGGTTGTCATACCTTGTATTTAGCAATTTA	495	
QY	772	TCTTGTTCCTTCAGTTTGGTATTATTATGATGCACAGAGCTTTCATGACATTAAGCC	831	
Db	496	TCTAATATTGTAAGTTTGGTATTATTATGATGCACAGAGCTTTCATGACATTAAGCC	555	
QY	832	TCTCTATAAGTATCTCCATGCCACCCATCATATATACAAAGCAAGATATACATCTCTCC	891	
Db	556	GCTTTACAAATATCTTCATGCTACCCATCATCTACAAATAAACAAGACACTCTCTCCCC	615	
QY	892	ATTTGCCGTAAGTGTTTTTCAGTTTGT	919	
Db	616	TTTGTGTTTGGCATTTTCCCTCTT	643	


```
; OTHER INFORMATION: any nucleotide
US-09-443-041A-29

Query Match      11.6%; Score 163.4; DB 4; Length 1210;
Best Local Similarity 70.6%; Pred. No. 2.1e-36;
Matches 218; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 592 AGATGCAATTCCTACAAATAAGGCTATGCGTTTGGCAATGTTGGCAATGAAGGCTAT 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 AGATGCTATCCCTACAGTAGAGCTATGAAGAAGCAATAATTTGTCATCAAGGCTAT 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 652 GCCATGGTACACTCTTCTCCAACTGCTCCGAGAGTAGTGAACGTGGTGGACCAA 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 GCCCTTCCTACTGCTCTCCGCTGCTATCTGAGCAGCATGATTGAGAGTGGATGGACAG 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 712 ATGTTTTCTAGCATAGAGCAATTCGGCTGCAATTCCTGATTTGTTTACATCGCCATCTA 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 GTGTTCTTCATATACAGGAAGTTGGTGGCTATGACATATCTATGTCCTCTATA 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 772 TCTCTATTCGTTGAGTTGGTATTATTTGATGACAGAGCTTCATGACATTAAGCC 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 TCTCATCTTTGTTGGAGTTCCGAATTTACTGGATGCACAGAGTTGTCATGACATAAGCC 427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 832 TCTCTATAGTATCTCCATGCGCCACCATCATATCTACAAACAGAGAGATACACTCTCTCC 891
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 ACTATACAGACCACTACATGTCACACCCACCACATTTTACAAACAGAGAGATACCCATCACC 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 892 ATTTGCCCG 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 ATTTGCTGG 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-443-041A-25
; Sequence 25, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 1318
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-443-041A-25

Query Match      11.5%; Score 161.8; DB 4; Length 1318;
Best Local Similarity 70.2%; Pred. No. 6.2e-36;
Matches 217; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 592 AGATGCAATTCCTACAAATAAGGCTATGCGTTTGGCAATGTTGGCAATGAAGGCTAT 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 351 AGATTCATACCTACAAATCGAAGCTATGAAGAAGCAATAATTTGTCATCAAGGCTAT 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 652 GCCATGTTACACTCTTCTCCAACTGCTCCGAGAGTAGTGAACGTGGTGGACCAA 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 411 GCCTCTCTATTGTCCTTCCAACTTATCTGAGTACATGTTGGTGAATGATGGACACA 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 712 ATGTTTCTAGCATAGAGCAATTCGCTGATTTCTGATTTGTTTACATCGCCATCTA 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 471 GTGTTATGTTATATACAGTAGAGTTGGTGGCAATGACCTGGTTTATCTGCGCTTATA 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 772 TCTTGTTCCTGTTGAGTTGGTATTATTTGATGACAGAGAGCTTCATGACATTAAGCC 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

;
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.7
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/978,687
; FILING DATE: FEBRUARY 1, 1993
; CLASSIFICATION: 800
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/05801
; FILING DATE: 15-AUGUST-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/568,493
; FILING DATE: 15-AUGUST-1990
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/721,761
; FILING DATE: 16-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; REFERENCE/DOCKET NUMBER: CGNE 76-2 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
;
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-07-978-687-31
;
; Query Match 2.6%; Score 37.2; DB 1; Length 2060;
; Best Local Similarity 46.2%; Pred. No. 0.73;
; Matches 123; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
;
; QY 365 GATTATTGTGCATATCTCCATCTAAGGGATTGAACAGTTAGTGGCTTATATAAGTTT 424
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 300 GATTAAAGTGGGAGATTGATTTTAAAGGATTTCGAAGATTTGAAGATTCATTGTTATT 241
;
; QY 425 TGTCCACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTCTACTTGCATTTAAGT 484
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 240 GCTTATATATCTTAAATTTTAAAAAGTCTTGTGTTATTGTTGTTGTTGTTGTTGTTAAT 181
;
; QY 485 CCACCTTAATTTGTTGTTGAAGTATTCTCTACATCTTTCAGACACATCTTTTCTGCTTCT 544
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 180 TCCATATAAATCTTTTGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 121
;
; QY 545 CTGAGACTCTGCTCTAGTTTGAATCTTTTGGTCTGTTTGGTCTGTTTGGTCTGTTTGGTCTGTT 604
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 120 ATTAATAAGCTGAGTATTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 61
;
; QY 605 ACAATAAGGCTATGCGTTTCAAAAT 630
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 60 TACACAAAAAATGTCACCTTTATAATT 35
;
; RESULT 15
; PCT-US91-01750-2/c
; Sequence 2, Application PC/TUS9101750
; GENERAL INFORMATION:
;
; APPLICANT: KNAUF, VIC C.
; APPLICANT: KRIDL, JEAN C.
; APPLICANT: SCHERER, DONNA E.
; TITLE OF INVENTION: Novel Sequences Preferentially
; TITLE OF INVENTION: Expressed in Early Seed
; TITLE OF INVENTION: Development and Methods
; TITLE OF INVENTION: Related Thereto
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01750
; FILING DATE: 19910314
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/494,722
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; REFERENCE/DOCKET NUMBER: CGNE 68WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 916-753-6313
; TELEFAX: 916-753-1510
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2060 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; PCT-US91-01750-2
;
; Query Match 2.6%; Score 37.2; DB 5; Length 2060;
; Best Local Similarity 46.2%; Pred. No. 0.73;
; Matches 123; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
;
; QY 365 GATTATTGTGCATATCTCCATCTAAGGGATTGAACAGTTAGTGGCTTATATAAGTTT 424
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 300 GATTAAAGTGGGAGATTGATTTTAAAGGATTTCGAAGATTTGAAGATTCATTGTTATT 241
;
; QY 425 TGTCCACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTCTACTTGCATTTAAGT 484
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 240 GCTTATATATCTTAAATTTTAAAAAGTCTTGTGTTATTGTTGTTGTTGTTGTTGTTAAT 181
;
; QY 485 CCACCTTAATTTGTTGTTGAAGTATTCTCTACATCTTTCAGACACATCTTTTCTGCTTCT 544
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 180 TCCATATAAATCTTTTGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 121
;
; QY 545 CTGAGACTCTGCTCTAGTTTGAATCTTTTGGTCTGTTTGGTCTGTTTGGTCTGTTTGGTCTGTT 604
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 120 ATTAATAAGCTGAGTATTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 61
;
; QY 605 ACAATAAGGCTATGCGTTTCAAAAT 630
; DB ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 60 TACACAAAAAATGTCACCTTTATAATT 35
;
; Search completed: June 6, 2003, 16:28:06
; Job time : 253.075 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 15:14:13 ; Search time 265.151 seconds
(without alignments)
7442.356 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552
Perfect score: 1410
Sequence: 1 atggcgccgataatgctta.....tgcacatcggaacatctg 1410

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870315 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	1410	100.0	1889	10	US-09-775-879-20	Sequence 20, Appl
2	587.4	41.7	2925	10	US-09-775-879-22	Sequence 22, Appl
3	309.4	21.3	846	9	US-09-938-842A-421	Sequence 421, Appl
4	125.4	8.3	285	10	US-09-294-093B-1568	Sequence 1568, Ap
5	106.2	7.5	304	10	US-09-294-093B-220	Sequence 220, Appl
6	73	5.2	277	10	US-09-923-876-5279	Sequence 5279, Ap
7	52.8	3.7	446	10	US-09-960-352-3400	Sequence 3400, Ap
8	52.6	3.7	529	9	US-09-983-965-2109	Sequence 2109, Ap
9	52	3.7	9539	9	US-10-239-676-52	Sequence 52, Appl
10	48.4	3.1	12465	9	US-10-239-676-31	Sequence 31, Appl
11	47.2	3.3	686	9	US-10-198-846-5499	Sequence 5499, Ap
12	47.2	3.3	6025	9	US-10-239-676-44	Sequence 44, Appl
13	46.8	3.3	2067	10	US-09-880-107-1750	Sequence 1750, Ap
14	46.8	3.3	2158	10	US-09-925-300-588	Sequence 588, Appl
15	46.8	3.3	10046	9	US-10-073-961-458	Sequence 458, Appl
16	46.8	3.3	10046	10	US-09-764-887-458	Sequence 458, Appl
17	46.2	3.3	8333	9	US-10-239-676-114	Sequence 114, Appl
18	45.8	3.2	9539	9	US-10-239-676-51	Sequence 51, Appl
19	45.2	3.2	425	10	US-09-960-352-4010	Sequence 4010, Ap

20	44.6	3.2	5690	9	US-10-239-676-73	Sequence 73, Appl
21	44.6	3.2	6544	9	US-10-239-676-46	Sequence 46, Appl
c	44.4	3.1	1267	12	US-10-001-843-45	Sequence 45, Appl
23	44	3.1	516	10	US-09-960-352-5785	Sequence 5785, Ap
24	44	3.1	7657	9	US-10-239-676-185	Sequence 185, Appl
25	43.8	3.1	11812	9	US-10-239-676-210	Sequence 210, Appl
26	43.6	3.1	12405	9	US-10-239-676-35	Sequence 35, Appl
27	43.4	3.1	640681	10	US-09-790-988-1	Sequence 1, Appl
28	43.2	3.1	5898	9	US-10-239-676-180	Sequence 180, Appl
29	43	3.0	2000	9	US-09-938-842A-3497	Sequence 3497, Ap
c	42.8	3.0	432	9	US-09-918-995-13651	Sequence 13651, A
31	42.8	3.0	15732	9	US-10-239-676-95	Sequence 95, Appl
32	42.6	3.0	428	10	US-09-960-352-573	Sequence 573, Appl
33	42.6	3.0	6306	9	US-10-239-676-224	Sequence 224, Appl
34	42.6	3.0	11047	9	US-10-239-676-188	Sequence 188, Appl
35	42.4	3.0	7195	9	US-10-239-676-30	Sequence 30, Appl
36	42.4	3.0	8333	9	US-10-239-676-113	Sequence 113, Appl
c	42.2	3.0	1682	10	US-09-925-302-298	Sequence 298, Appl
38	42.2	3.0	6046	9	US-10-239-676-15	Sequence 15, Appl
39	42	3.0	393	10	US-09-960-352-4582	Sequence 4582, Ap
c	41.6	3.0	11735	9	US-10-239-676-34	Sequence 34, Appl
41	41.6	3.0	525	9	US-10-198-846-1483	Sequence 1483, Ap
42	41.4	2.9	639	10	US-09-878-574-4316	Sequence 4316, Ap
43	41.4	2.9	5127	9	US-10-239-676-132	Sequence 132, Appl
c	40.8	2.9	994	9	US-10-198-846-6968	Sequence 6968, Ap
45	40.6	2.9	5337	9	US-10-239-676-137	Sequence 137, Appl

ALIGNMENTS

RESULT 1
US-09-775-879-20
; Sequence 20, Application US/09775879
; Patent No. US20020088822A1
; GENERAL INFORMATION:
; APPLICANT: Choe, Sunghwa
; APPLICANT: Feldmann, Kenneth A.
; TITLE OF INVENTION: Dwf7 MUTANTS
; FILE REFERENCE: 2225-0003
; CURRENT APPLICATION NUMBER: US/09/775,879
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/179,901
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; TYPE: DNA
; ORGANISM: Genomic dwf7 (Arabidopsis)
US-09-775-879-20

Query Match	100.0%	Score 1410	DB 10	Length 1889
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1410	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGCGCGCGGATAATGCTTATCTGATGCGAGTGTGACGAAACCTCTTTTACACACCA	60	
Db	143	ATGCGCGCGGATAATGCTTATCTGATGCGAGTGTGACGAAACCTCTTTTACACACCA	202	
QY	61	ATCGTTCGAGTCATCTTTTTCGCGGGAATCTATGGAAACCCCTACCTCATTTTCTCCAG	120	
Db	203	ATCGTTCGAGTCATCTTTTTCGCGGGAATCTATGGAAACCCCTACCTCATTTTCTCCAG	262	
QY	121	ACATGGCTCCGAAATACCTCCCGGAACCTACTATATCTCTCCGGTTCTCTCTGG	180	
Db	263	ACATGGCTCCGAAATACCTCCCGGAACCTACTATATCTCTCTCCGGTTCTCTCTGG	322	
QY	181	TGCTTCTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAAGTCTCGACTTTCAC	240	
Db	323	TGCTTCTACATCTATTACCTTAAATCAACGTTTACCTTCCCAAAGTCTCGACTTTCAC	382	
QY	241	TTTGTATTCACTATTGCTTAATCGCTTTCTATGTTATCGATTTTCAATTTAAGGAAGA	300	

Db 383 |||||TTTGTATTACCTATTGCTTAATCGCTTCTATGTTATCGATTTTTCATTTTAAAGGAAGA 442
QY 301 GGGTTTCTTCGCTGCTACTGTACAGTAATTTGGATTTGATGCTGGATAGTCTCATGTTTGCATTT 360
Db 443 GGGTTTCTTCGCTGCTACTGTACAGTAATTTGGATTTGATGCTGGATAGTCTCATGTTTGCATTT 502
QY 361 TATTGATTATTGTTGTCATATTTCCATCTAAGGATTGAACAGTTAGTGGCTTATATAAG 420
Db 503 TATTGATTATTGTTGTCATATTTCCATCTAAGGATTGAACAGTTAGTGGCTTATATAAG 562
QY 421 TTTTGTGCAACCAATGAGAAGTCGACATCTTTTGAAGTTGAATTTTCTACTTGCATTT 480
Db 563 TTTTGTGCAACCAATGAGAAGTCGACATCTTTTGAAGTTGAATTTTCTACTTGCATTT 622
QY 481 AAGTCCACCTAAATCTTTGTTGAAGTATGCTACTTTCACACACATCTTTTCTGC 540
Db 623 AAGTCCACCTAAATCTTTGTTGAAGTATGCTACTTTCACACACATCTTTTCTGC 682
QY 541 TTCTCTGAGACTCTGCTTAGTTTGAATCTTTTGGTCTGTTTGGCTTCAGATCAAT 600
Db 683 TTCTCTGAGACTCTGCTTAGTTTGAATCTTTTGGTCTGTTTGGCTTCAGATCAAT 742
QY 601 TCTCAATAAAGGCTATGCGTTTGCATATGTTTGGCAATGTTTGGCAATGAGGCTATGCCATGTA 660
Db 743 TCTCAATAAAGGCTATGCGTTTGCATATGTTTGGCAATGTTTGGCAATGAGGCTATGCCATGTA 802
QY 661 CACTCTCTTCCAACTGCTCCGAGAGTATGATGAACGTGTTGGACCAAAATGTTTTCG 720
Db 803 CACTCTCTTCCAACTGCTCCGAGAGTATGATGAACGTGTTGGACCAAAATGTTTTCG 862
QY 721 TAGCATAGACGAATTCGGCTGGATCTGTATTTTGTATACATCGCCATCTATCTTGT 780
Db 863 TAGCATAGACGAATTCGGCTGGATCTGTATTTTGTATACATCGCCATCTATCTTGT 922
QY 781 CGTTGAGTTGGTATTTTGTATGATGACAGAGCTTTCATGACATTAAGCCCTCTCTATAA 840
Db 923 CGTTGAGTTGGTATTTTGTATGATGACAGAGCTTTCATGACATTAAGCCCTCTCTATAA 982
QY 841 GTATCTCCATGCCACCCATCATATCFACAACAGAGAGTATACACTCTCTCCATTTGCCGG 900
Db 983 GTATCTCCATGCCACCCATCATATCFACAACAGAGAGTATACACTCTCTCCATTTGCCGG 1042
QY 901 TAAGTGTTCAGTTTGTCTTCTTTAGTTCTTGTAAAGATTGGTAGCATTTAGTTTCT 960
Db 1043 TAAGTGTTCAGTTTGTCTTCTTCTTTAGTTCTTGTAAAGATTGGTAGCATTTAGTTTCT 1102
QY 961 TACCAGAAAGACTTTGTGACGAGCTGCTGTACTCCAAATCAGATTTTGCATTTCCCTTAT 1020
Db 1103 TACCAGAAAGACTTTGTGACGAGCTGCTGTACTCCAAATCAGATTTTGCATTTCCCTTAT 1162
QY 1021 CCATAAAGTAAACAGAAAGCTAGATTTATAAATGTACGTGCAATTTACTTCCATATG 1080
Db 1163 CCATAAAGTAAACAGAAAGCTAGATTTATAAATGTACGTGCAATTTACTTCCATATG 1222
QY 1081 TCAGAGAGACTTCTGACTTAAACAGAGTTAGATCTTTGTGTTCTTCTGCTCGGA 1140
Db 1223 TCAGAGAGACTTCTGACTTAAACAGAGTTAGATCTTTGTGTTCTTCTGCTCGGA 1282
QY 1141 CTGATTGGAATGACGAGAGTTCTTTTATCTTCTCCCTGGAGTGTATCTGGTTAATC 1200
Db 1283 CTGATTGGAATGACGAGAGTTCTTTTATCTTCTCCCTGGAGTGTATCTGGTTAATC 1342
QY 1201 CAAGGATGTGACATCTAATATTACTTGTAACTTCTTACGTTTGTTCACAGGGCTTGC 1260
Db 1343 CAAGGATGTGACATCTAATATTACTTGTAACTTCTTACGTTTGTTCACAGGGCTTGC 1402
QY 1261 ATTTCACCGATAGAGGGATCTTTCAGGCTGTACCGATGATGATAGCGCTGTTTATAGT 1320
Db 1403 ATTTCACCGATAGAGGGATCTTTCAGGCTGTACCGATGATGATAGCGCTGTTTATAGT 1462
QY 1321 GCCAATTTCATTCACAACTCATATAGTCTTTTGTTCATGGAAGCGATATGACGGCGAA 1380
|||||

Db 1463 GCCAATTTCATTCACAACCTCATATAGTCTTTTGTTCATGGAAGCGATATGACGGCGAA 1522
QY 1381 CATCCATGACTGCATCCATGGCAACATCTG 1410
Db 1523 CATCCATGACTGCATCCATGGCAACATCTG 1552

RESULT 2
US-09-775-879-22
; Sequence 22, Application US/09775879
; Patent No. US20020068822A1
; GENERAL INFORMATION:
; APPLICANT: Choe, Sunghwa
; APPLICANT: Feldmann, Kenneth A
; TITLE OF INVENTION: Dwf7 MUTANTS
; FILE REFERENCE: 2225-0003
; CURRENT APPLICATION NUMBER: US/09/775,879
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/179,901
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 2925
; TYPE: DNA
; ORGANISM: Genomic HDF7
US-09-775-879-22

Query Match 41.7%; Score 587.4; DB 10; Length 2925;
Best Local Similarity 99.4%; Pred. No. 4e-142; 1; Indels 3; Gaps 3;
Matches 621; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 787 GTTTGGTATTTATTGGATGCACAGAGCTTTCATGACATTAAGCCCTCTCTATAAGTATCT 846
Db 1 GTTGGTATTTATTGGATGCACAGAGCTTTCATGACATTAAGCCCTCTCTATAAGTATCT 60
QY 847 CCATGCCACCCATCATATCTACAACAGAGAGTATACACTCTCTCCATTTGCCGGTAAGTG 906
Db 61 CCATGCCACCCATCATATCTACAACAGAGAGTATACACTCTCTCCATTTGCCGGTAAGTG 120
QY 907 TTTTCAGTTTGTCTTCTTTAGTTCTTGTAAAGATTGGTAGCATTTAGTTTCTTTACCAG 966
Db 121 TTTTCAGTTTGTCTTCTTTAGTTCTTGTAAAGATTGGTAGCATTTAGTTTCTTTACCAG 180
QY 967 AAAAGACTTTGTGACGAGCTGCTGTACTCCAAATCAGATTTTGCATTTCCATCCATAA 1026
Db 181 AAAAGACTTTGTGACGAGCTGCTGTACTCCAAATCAGATTTTGCATTTCCATCCATAA 240
QY 1027 AGTAACCAAGAGGCTAGAAATATATAATGTCAGCTGCAATTTACATATGTCAGAG 1086
Db 241 AGTAACCAAGAGGCTAGAAATATATAATGTCAGCTGCAATTTACATATGTCAGAG 300
QY 1087 AGACTTCTCAGTTAAACAGAGTTAGATCTTTGTGTTTCTCTCTCGGACTGATT 1146
Db 301 AGACTTCTCAGTTAAACAGAGTTAGATCTTTGTGTTTCTCTCTCGGACTGATT 360
QY 1147 GGAAATGACGAGAAGTTCTTTTATCTACTTCCCTGGAGTGTATCTTGGTTAATCCAAGGA 1206
Db 361 GGAAATGACGAGAAGTTCTTTTATCTACTTCCCTGGAGTGTATCTTGGTTAATCCAAGGA 420
QY 1207 TGTGACATCT - AATATTACTTGTAACTTCCCTAGTTTGTTCAGGGCTTGCATTTTC 1265
Db 421 TGTGACATCTAATATTTACTTGTAACTTCCCTAGTTTGTTCAGGGCTTGTGA - TTC 479
QY 1266 ACCCAGTACAGGGATPACTTACGGCTGTACCGCATGTAGAGGCTGTTTATAGTCCCAA 1325
Db 480 ACCCAGTACAGGGATPACTTACGGCTGTACCGCATGTAGAGGCTG - TTATAGTCCCAA 538
QY 1326 TTCAATTCACAACCTCATATAGTCTTTTGTTCATGGAAGCGATATGACGGCGGAATCC 1385
Db 539 TTCAATTCACAACCTCATATAGTCTTTTGTTCATGGAAGCGATATGACGGCGGAATCC 598
QY 1386 ATGACTGCAATCCATGCCAACATCTG 1410


```
QY 735 TCGGCTGGATTCGTATTTTGTATACGCGCATCTATCTCTTTTCGGTTGAGTTGGTA 794
Db 61 TTGGTTTTCTATGACCTCTGTATATATGCTATGCTATCTATCTTTTGGAGTTGGAA 120
QY 795 TTTATTTGGATGACAGAGCTTCATGACATTAAAGCCCTCTCTAAGTATCTCCATGCCA 854
Db 121 TTTACTGGATGACAGAGAGTTGTCATGACATAAAACCAATATACAAATATCTGCATGCAA 180
QY 855 CCCATCATATCTACACAGACAGATACACTCTCTCCATTTTCCCGTAAAGTCTTT 909
Db 181 CCCACATATTTACACAGAGGAATAC-TTGTCTCCATTTCTGCTGGACTGCATTT 234

RESULT 6
US-09-923-876-5279
; Sequence 5279, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalguodi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Shetman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5279
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700456730H1
US-09-923-876-5279
```

```
Query Match 5.2%; Score 73; DB 10; Length 277;
Best Local Similarity 65.1%; Pred. No. 4e-09; Mismatches 65; Indels 1; Gaps 1;
Matches 123; Conservative 0;

QY 592 AGATGCAATTCCTACAATAAAGGCTATGCGTTTGCAAAATGTTTGCGCAATGAAGGCTAT 651
Db 84 AGATGCCATCCCAACAAATGAAGCTATGAAGAAGCAAAATAGCTAGCATCCCAAGGCTAT 143
QY 652 GCGATGGTACACTCTTCTCCAACTGTCTCCGAGAGTATGATGAACGTGGTTGGACCAA 711
Db 144 GCGTTTTACTGTCTCTTCCAACTTTATCTGAGTATATGATCGAGAGCGGATGGACCG 203
QY 712 ATGTTTGTACATAGAGCAATTCGGCTGGATTCTGTATTTTGTATACATCGCCATCTA 771
Db 204 GTGTTACTTTATATACAGCAAAATGGTTTTTCTCGGTACCTTG-TTATATGGCTATGTA 262
QY 772 TCTTGTTTT 780
Db 263 TCTCATTTT 271
```

```
RESULT 7
US-09-960-352-3400
; Sequence 3400, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
```

```
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3400
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 15-LIB3058-052-Q1-K1-D11
US-09-960-352-3400

Query Match 3.7%; Score 52.8; DB 10; Length 446;
Best Local Similarity 46.9%; Pred. No. 0.00092; Mismatches 187; Indels 0; Gaps 0;
Matches 165; Conservative 0;

QY 235 TTTCACTTTTGTATTCACATATTCGTAATCGCTTTCTATGTTATCGATTTTCAATTTAA 294
Db 95 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 154
QY 295 GGAAGAGGGTTTCTTCGTACTGTACAGATAATTTGGATTTTGATGTGGATATGTTTCATGTT 354
Db 155 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 214
QY 355 TGCATTTATGATTTTGTGCAATATCTCCATCTAAGGATTAACAGTTAGTGCCTTA 414
Db 215 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 274
QY 415 TATAAGTTTCTGCAACCAATGAGAAGTCGTACATCTTTGAAGTTGAATTTTCTACTTG 474
Db 275 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 334
QY 475 CCATTTAAGTCACATTAATTTGTTGAACTGATTTGCTACTTTTCAGACACATCTTTT 534
Db 335 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 394
QY 535 TTTCTGCTCTCTGAGACCTGCTTAGTTTGAATCTTTTGGTCTGTTTT 586
Db 395 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 446
```

RESULT 8

```
US-09-983-965-2109
; Sequence 2109, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2109
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (335)
; OTHER INFORMATION:
; OTHER INFORMATION: Clone ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109
```

```
Query Match 3.7%; Score 52.6; DB 10; Length 529;
Best Local Similarity 46.6%; Pred. No. 0.0011; Mismatches 190; Indels 0; Gaps 0;
Matches 166; Conservative 0;
```



```
Db 2017 TTTTATGATGATTTTATAGTTAAAT--ATAGAAATAGAGTAGTTTTTTTTTTGA 2074
QY 523 GACACATCTTTTCTGCTCTGAGACTCTGCTAGTCTTGAATCTTTTGTGCTG 582
Db 2075 AATATATGTTTTTATTTTATTTTATTTTATTTTATTTTATTTATTTAT 2134
QY 583 TTTTGTTCAGATGCAATTCCTACAATAAGGCTATGCGTTT 624
Db 2135 TTTTGTAAATGAAAGGTAGAAAGTAAATTTGTTGCT 2176

RESULT 11
US-10-198-846-5499
; Sequence 5499, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2, 4, 22, 33, 44, 149, 196, 204, 205, 206, 210, 213, 220,
; LOCATION: 227, 228, 229, 230, 231, 244, 247, 248, 249, 250, 251, 252,
; LOCATION: 253, 254, 255, 256, 261, 271, 272, 276, 277, 279, 282, 283,
; LOCATION: 284, 285, 286, 291, 292, 294, 295, 297, 298, 299, 302
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314,
; LOCATION: 315, 317, 318, 319, 320, 321, 328, 334, 336, 343, 344,
; LOCATION: 347, 348, 349, 350, 351, 353, 355, 359, 360, 362, 363, 368,
; LOCATION: 369, 370, 371, 372, 373, 379, 384, 385, 387, 391, 392
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 396, 397, 398, 399, 402, 405, 406, 407, 408, 413, 415,
; LOCATION: 416, 426, 430, 433, 434, 438, 440, 441, 442, 445, 448, 463,
; LOCATION: 468, 469, 471, 486, 487, 488, 489, 490, 491, 493, 494, 496,
; LOCATION: 497, 499, 500, 501, 502, 505, 510, 511, 515, 517, 518
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 520, 527, 528, 529, 535, 536, 539, 557, 559, 562, 565, 567,
; LOCATION: 569, 570, 573, 581, 582, 584, 585, 590, 599, 602, 604, 608,
; LOCATION: 616, 617, 619, 623, 626, 627, 630, 632, 633, 634, 635, 636,
; LOCATION: 637, 638, 641, 643, 644, 648, 651, 653, 655, 658, 661
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 662, 665, 669, 676, 677
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5499
Query Match 3.3%; Score 47.2; DB 9; Length 686;
Best Local Similarity 31.4%; Pred. No. 0.033;
Matches 109; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 243 TTGATTCATATGCTTAATCGCTTCTCTATGTTATCGATTTTCAATTAAGGAGG 302
```

```
Db 257 TTTTNTCTCTTTTNTTNNCTTNNNNCTTNNNNCTTNNNNCTNNNNNNNNNNNT 316
QY 303 GTTCTCTCTGCTACGCTACAGTAATTTGGATTGATGCTGATAGTTCATGTTGCATTTA 362
Db 317 NNNNNNTTTTNTTNTTNCNTCTTTTNTTNNNNNTTNTTNTTNNNNCTTNNNNNNNTT 376
QY 363 TTGATTATTGTCATATTCTCCATCTAAGGATTGAACAGTAGTGGCTTATATAAGTT 422
Db 377 TTTTNTTNTTNTTNNNTTNNNNNTTNTTNNNNNTTNTTNTTNTTNTTNTTNTTNTT 436
QY 423 TTTGTGCAACCAATGAGAGTGTGATCATCTTTTGAAGTTGAAATTTTCTACTTGCATTTAA 482
Db 437 TTTNNNNCTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 496
QY 483 GTCCACTTAAATTTGTTGTTGAAGTGTCTACTTTTCAGACACATCTTTTCTGCTT 542
Db 497 NTNNNNNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 556
QY 543 CTCTGAGACTCTGCTCTAGTTTGAATCTTTTGTGCTGTGTTTGTCT 589
Db 557 NTNTTNTTNTTNNCTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 603

RESULT 12
US-10-239-676-44
; Sequence 44, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 44
; LENGTH: 6025
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-239-676-44
Query Match 3.3%; Score 47.2; DB 9; Length 6025;
Best Local Similarity 45.8%; Pred. No. 0.11;
Matches 163; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 235 TTTTCTTTTGTATTCACACTATTTGCTTAATCGCTTTCTATGTTATCGATTTTCAATTTAA 294
Db 2987 TTTTCTTTTGTATTTAGTTTGTGTTGGCGATATAGTGAATTTTAAAAAATAAAAAA 3046
QY 295 GGAAGAGGGTTTCTTCTGCTACTGTACACATAATTTGGATTGATGGGATAGTTCATGTT 354
Db 3047 AAATGATTAATAAACGGATAAGTGGATTTTATATAGAGATTATTTATGTTATTAG 3106
QY 355 TGCATTTATGATTTATTTTGTGCATATCTCCATCTTAAGGATTAAGACAGTGTAGTGGCTTA 414
Db 3107 ATTATTAAATAATTTTCTGTCGCTGTTTATTTTGTGTTAGAGTGGTGTGA 3166
QY 415 TATAAGTTTCTGTCGAACCAATGAGAGTCGTACATCTTTTGAAGTTGAATTTTCTACTGTT 474
```

Db 3167 GGTATTTTTCAGAGAAAAAGATAGATAGATCTTTTATTTTTTTTTTTT 3226
Qy 475 CCATTTLAGTCCACTTAAATGTTTGTGAAGTGTCTTACCTTTCAGACACATCTTT 534
Db 3227 TTTTTLTT 3286
Qy 535 TTTGCTCTCTGTGAGACTCTCTTCTAGTTTGAATCTTTTGTGCTGTTTGTCTT 590
Db 3287 TTTTTLTT 3342

RESULT 13

US-09-880-107-1750
; Sequence 1750, Application US/09880107
; Patent No. US2002042981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scher, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1750
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D85181
US-09-880-107-1750

Query Match 3.3%; Score 46.8; DB 10; Length 2067;
Best Local Similarity 57.5%; Pred. No. 0.078;
Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy 1247 TTTACAGGGCTTGCACTTACCCAGTAGAGGGGATCTTACGGCTGTACCGCATGTGATA 1306
Db 747 TTTGCAAGTCATGCTTTTACCCCTATTGATGGCTTCTTTCAGAGCTTACCTTACCATA 806
Qy 1307 GCGCTGTTTATAGTCCCAATTCATTTCAACACTCATATAGTCTTTTGTTCATGGAAGCG 1366
Db 807 TACCTTTTATCTTCCATTACACAAGGTGGTTTATTAAGTCTGTACATCTTGGTTAAT 866
Qy 1367 ATATGGACGGGCAACATCCATGACTG 1392
Db 867 ATCTGGACAATTTCCATTTCATGACGG 892

RESULT 14

US-09-925-300-588
; Sequence 588, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 588

; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-588
Query Match 3.3%; Score 46.8; DB 10; Length 2158;
Best Local Similarity 57.5%; Pred. No. 0.08;
Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
Qy 1247 TTTACAGGGCTTGCACTTACCCAGTAGAGGGGATCTTACGGCTGTACCGCATGTGATA 1306
Db 596 TTTGCAAGTCATGCTTTTACCCCTATTGATGGCTTCTTTCAGAGCTTACCTTACCATA 655
Qy 1307 GCGCTGTTTATAGTCCCAATTCATTTCAACACTCATATAGTCTTTTGTTCATGGAAGCG 1366
Db 656 TACCTTTTATCTTCCATTACACAAGGTGGTTTATTAAGTCTGTACATCTTGGTTAAT 715
Qy 1367 ATATGGACGGGCAACATCCATGACTG 1392
Db 716 ATCTGGACAATTTCCATTTCATGACGG 741

RESULT 15

US-10-073-961-458
; Sequence 458, Application US/10073961
; Publication No. US20030077602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924

;
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17

;
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221

; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,475
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/231,243
 ; PRIOR FILING DATE: 2000-09-08

Query Match	3.3%	Score 46.8;	DB 9;	Length 10046;
Best Local Similarity	57.5%	Pred. No. 0.19;		
Matches 84;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

Qy	1247	TTTACACGGCTTGCATTTACCCAGTAGACACGGGATACCTTCAGGCTGTACCGCATGTGATA	1306
Db	3736	TTTGCA/STCATGCTTTTCACCCCTATTGATGGCTTTCTTCAGAGTCTACCTTACCATATA	3795
Qy	1307	GCCTGT/TTATAGTGCCAAATTCATTTCACAACCTCATATAGTCTTTTGTTCATGGAACG	1366
Db	3796	TACCTT/TTATCTTTCCATTACACAAGGTGTTTATTAAAGTCGTACATCTTGGTTAAT	3855
Qy	1367	ATATGG/CGCGGAACATCCATGACTG	1392
Db	3856	ATCTGG/CAATTTCCATTCATGACGG	3881

Search completed: June 6, 2003, 18:14:19
 Job time : 271.151 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 12:27:48 ; Search time 3665.11 Seconds
(without alignments)
6230.541 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_1552

Perfect score: 1410

Sequence: 1 atigcgcgcgataatgctta.....tgcatccatgcgaacatctg 1410

Scoring table: IDENTITY NUC
Gap: 10.0 , Gapext 1.0

Searched: 16151066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: an_estba.*

2: an_esthum.*

3: an_estin.*

4: an_estnu.*

5: an_estov.*

6: an_estpl.*

7: an_estro.*

8: an_hic.*

9: jb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_man.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	478	33.9	498	17	AQ011450 F27K17TFC
2	183.6	13.0	583	12	BF425767 sr42f04.y
3	182	12.9	716	10	AW775288 EST334353
4	182	12.9	804	13	BI308419 EST529829
5	181.8	12.9	535	10	AW596303 sj01f11.y
6	181.4	12.9	592	10	AW219366 EST301848

7	173.8	12.3	667	10	BE357415
8	173.4	12.3	621	12	BE123510
9	173.4	12.3	753	13	BI930355
10	170.2	12.1	722	12	EG598293
11	169.6	12.0	574	14	Q803827
12	168.8	12.0	708	10	BE602261
13	168.2	11.9	612	14	Q472907
14	168.2	11.9	769	14	Q802075
15	168	11.9	597	12	EG098248
16	167.2	11.9	708	10	BE601702
17	166.8	11.8	608	14	Q118369
18	166.6	11.8	782	12	BF624091
19	166.6	11.8	782	12	BF624091
20	166.6	11.8	782	12	BF624091
21	166	11.8	786	14	Q806050
22	165	11.7	668	10	BE404301
23	165	11.7	781	12	BF617574
24	163.4	11.6	658	14	Q246840
25	161.8	11.5	605	13	BI417093
26	161.8	11.5	605	13	BI417093
27	158.2	11.2	517	14	Q703739
28	158	11.2	570	13	EG933493
29	157.2	11.1	652	14	Q766124
30	155	11.0	587	12	BF586850
31	152.8	10.8	586	14	BM903391
32	150.6	10.7	984	11	AY110327
33	146.8	10.4	667	17	BH424776
34	145.8	10.3	661	10	BE413481
35	145.2	10.3	562	10	BE125051
36	145	10.3	567	10	AW429266
37	141	10.0	691	13	BI958272
38	135.4	9.6	840	12	EG320935
39	134.4	9.5	486	13	BI643821
40	132.6	9.4	583	14	Q695873
41	124.8	8.9	541	12	EG322635
42	124.6	8.8	572	10	BE471514
43	120	8.5	586	14	BU012147
44	118.8	8.4	551	10	BE471892
45	116.6	8.3	464	10	AW040702

ALIGNMENTS

RESULT 1
AQ011450/c
LOCUS F27K17TFC IGF Arabidopsis thaliana genomic clone F27K17, DNA
DEFINITION 498 bp DNA linear GSS 29-MAY-1998
sequence.
ACCESSION AQ011450
VERSION AQ011450.1 GI:3166695
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Arabidopsis thaliana
AUTHORS Rounsley S.D., Suh E.J., Wible C., Golden K., Shatsman S., Choi P.,
Yu K., Akinretaye B., Shen K., Goonasekaram S., Millscher J.,
Adams M.D. and Venter J.C.
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 4
JOURNAL Unpublished (1998)
COMMENT Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends

```
FEATURES
  source
    High quality sequence stop: 498.
    Location/Qualifiers
      1..498
        /organism="Arabidopsis thaliana"
        /strain="Columbia"
        /db_xref="taxon:3702"
        /clone_lib="F27K17"
        /sex="hermaphrodite"
        /note="Vector: Bel08A11; Site_1: EcoRI; Site_2: EcoRI;
        Produced by Thomas Altmann"
      200 a 95 c 83 g 120 t

BASE COUNT
ORIGIN
  Query Match 33.9%; Score 478; DB 17; Length 498;
  Best Local Similarity 98.8%; Pred. No. 6.5e-99;
  Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 221 CCAAAGGCTCGACTTTCATTTGATTTCACATATTGCTTAATCGCTTTCTATGTTATCG 280
DB 498 CCAAAGGCTCGACTTTCATTTGATTTCACATATTGCTTAATCGCTTTCTATGTTATCG 439
QY 281 ATTTTTCATTTT-AAGGAAGAGGGTTTCTCGTGACTGTACAGTAATTTGGATTGATG 339
DB 438 ATTTTTCATTTTAAACAGAGGGTTTCTCGTGACTGTACAGTAATTTGGATTGATG 379
QY 340 TGGATAGTTCATGTTTGCATTTATTTGATTATTTGTCATATTTCCATCTAAGGGATTGA 399
DB 378 TGAATAGTTCATGTTTGCATTTATTTGATTATTTGTCATATTTCCATCTAAGGGATTGA 319
QY 400 ACAGTTAGTGGCTATATAAGTTTGTGCAACCAATGAGAGTCTACATCTTGAAGT 459
DB 318 ACAGTTAGTGGCTATATAAGTTTGTGCAACCAATGAGAGTCTACATCTTGAAGT 259
QY 460 TGAATTTTCTACTTGCATTTAAGTCCACTTAATTTGTTTGAAGTGAATGCTCTACTT 519
DB 258 TGAATTTTCTACTTGCATTTAAGTCCACTTAATTTGTTTGAAGTGAATGCTCTACTT 199
QY 520 TCAGACACATCTTTTCTCTCTGAGACTCTCTCTAGTTTGAATCTTTTGTGT 579
DB 198 TCAGACACATCTTTTCTCTCTGAGACTCTCTCTAGTTTGAATCTTTTGTGT 139
QY 580 CTGTTTGTCTGAGTGCATTCCTACATAAAGGCTATGCGTTTCCAATGTTTGTGCG 639
DB 138 CTGTTTGTCTGAGTGCATTCCTACATAAAGGCTATGCGTTTCCAATGTTTGTGCG 79
QY 640 AATGAAGGCTATGCCATGTTACACTCTCTTCCAACCTGCTCCGAGAGTATGATCAACG 699
DB 78 AATGAAGGCTATGCCATGTTACACTCTCTTCCAACCTGCTCCGAGAGTATGATCAACG 19
QY 700 TGGTTGGACCAATGTTT 717
DB 18 TGGTTGGACCAATGTTT 1

RESULT 2
BF425767 583 bp mRNA linear EST 06-DEC-2001
LOCUS sr42f04.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl051-728 5', similar to TR:Q9Z129 Q9Z129 STEROL-C5(6
)-DESATURASE. ;, mRNA sequence.
ACCESSION BF425767
VERSION BF425767.1 GI:11413756
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 583)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 393.
Location/Qualifiers
  1..383
    /organism="Glycine max"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl051-728"
    /clone_lib="Gm-cl051"
    /tissue_type="floral meristematic mRNA"
    /lab_host="DH10B"
    /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
    XhoI; The cDNA library was constructed from floral
    meristematic mRNA provided by Dr. Halina Knap of Clemson
    University. Complementary DNA was synthesized from mRNA
    using a primer consisting of a poly(dT) sequence with a
    XhoI restriction site. EcoRI adapters were ligated to the
    blunt-ended cDNA fragments followed by XhoI digestion. The
    cDNA fragments were directionally cloned into the
    EcoRI-XhoI restriction site of the pBluescript vector. The
    ligated cDNA fragments were transformed into DH10B host
    cells (GibcoBRL). This library was constructed in the
    laboratory of Dr. Randy Shoemaker."
  131 a 149 c 120 g 183 t

BASE COUNT
ORIGIN
  Query Match 13.0%; Score 183.6; DB 12; Length 583;
  Best Local Similarity 71.1%; Pred. No. 1.4e-31;
  Matches 243; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 578 GTCTGTTTGTCTTCAGATGCAATTCCTACAAATGAAGGCTATGCGTTTGCAAATGTTGTG 637
DB 203 GTTATGTTCCCAAGATGCTATTCCTCTCGAAGAGCCATGCTTTGCAAAATATCTGTT 262
QY 638 GCATGAAGGCTATGCCATGTTACACTCTTCTTCCAACCTGCTCCGAGAGTATGATTGAA 697
DB 263 GCTATGAAGCCATGCTGTTGGTACACTTTCCTTCCAACCTGTTTCGGAGTACCTGTAGAA 322
QY 698 CGTGGTTGGACCAAAATGTTTGTCTAGCATAGAGAAATTCGGCTGGAATCTGTATTTGTT 757
DB 323 ACTGGCTGGACAAAGTCTATCTCTAGATATATAATGTTGGTGGCTTGCATACCTTGTG 382
QY 758 TACATGCCATCTATCTTGTGTTTTCGTTGAGTTTGGTATTTATTTGGATGCACAGAGCTT 817
DB 383 TATTTAGCAATTTATCTAAGTATTTAGAGTTTGGTATTTATTTGGATGCACAGAGACTG 442
QY 818 CATGACATTAAGCCTCTCTATAGTATCTCCATGCCACCCATCATATCTACACACAGCAG 877
DB 443 CACGACATAAAACCCGCTTTTACAAATATCTTCTGCTACCCATCCACATACATACAAACAG 502
QY 878 AATACACTCTCCCATTTTCGCGTAAGTGTCTTTCAGTTTGT 919
DB 503 AACACTCTCTCCCTTGTGCTGTTGGCATTTTCATCTCT 544

RESULT 3
AW775288
LOCUS AW775288
DEFINITION EST334353 DSIL Medicago truncatula cDNA clone pDSIL-1E1, mRNA
716 bp mRNA linear EST 07-SEP-2000
EST334353
```

sequence.
 accession AW75288 1 GI:7765101
 version EST.
 keywords barrel medic.
 organism Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases: 1 to 716)
 Pedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S., Peng
 H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S.,
 Holt, I.E. and Fraser, C.M.
 ESTs from leaves of Medicago truncatula after inoculation with
 Colletotrichum trifolii
 Unpublished (2000).
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlugh Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: dabbys@puccini.crl.umn.edu
 Minnesota; sequence name: M258621e
 TIGR sequence name: MTFAA25TK
 More information is available at:
 http://cirsyle.tamu.edu/medicago
 Seq primer: SKmod (CTA gaa cta gta gAT CC).
 Location/Qualifiers
 1..716
 /organism="Medicago truncatula"
 /cultivar="genotype Al7"
 /db_xref="taxon:3880"
 /clone="pDSIL-1E1"
 /clone_lib="DSIL"
 /tissue_type="leaves infected with Colletotrichum
 trifolii"
 /dev_stage="cotyledons and primary leaves harvested 5 and
 8 days after inoculation with Colletotrichum trifolii"
 /lab_host="E. coli strain XL0R"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 cotyledons and primary leaves harvested 5 and 8 days after
 inoculation with Colletotrichum trifolii. The cDNA was
 directionally ligated into the Uni-ZAP XR vector from
 Strategene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-ZAP phage using Ex-Assist
 helper phage and propagated in XL0R cells. Note: EST may
 be of fungal origin."
 187 a 162 c 112 g 255 t

FEATURES source

Query Match 12.9%; Score 182; DB 10; Length 716;
 Best Local Similarity 70.8%; Pred. No. 3.3e-31;
 Matches 242; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 578 GTCTGTTTCTTCAGATGCAATTCCTACATAAAGGCTATGCGTTTGCAAAATGTTGTG 637
 296 GTTTAT TCTCCATAAAGATTCATCCATCAGTAGAGCCATGTTCTTGCAAAATATCGTT 355
 638 GCAATGAAGGCTATGCCATGTACACTCTCTCCAACTGCTCCGAGAGTATCATGAA 697
 356 ACATGAAGGAATGCCTTGGTATAGTTGTTCCCACTATTTCTGAGTGCTCGGTAGAA 415
 698 CGTGGT TGGACCAAAATGTTTGTAGCATACAGAAATTCGGCTGATTTCTGATTTGTT 757
 416 CTTGGCTGGACAAAATCCTTCCCTAGATAAATAATCTTGGTTGGATTCCTACATTTGT 475
 758 TACATC TCCATCATCTGTTTTCGTTGAGTTTGGTATTAATGAGTCACAGAGAGCTT 817
 476 TATTAT TATATATGTTTATGAGAGTTTGGTATTTATTTGATGATGATGATGATG 535

QY 818 CATGACATTAAGCCTCTCTATAAGTATCTTCATGCCATCATATCTACAAACAGCAG 877
 536 CATGACATAAACCCTTTACAAATATCTTCATGCAACACATCATCTACATAACAA 595
 QY 878 AATACATCTCTCCATTTGCGGTAAGTGTGTTTTCAGTTTGTGTT 919
 596 AATACATCTCTCCATTTGCTGGTTGGCTTTTCATCTCTCTT 637
 RESULT 4
 LOCUS BI308419
 DEFINITION EST529829 GP0D Medicago truncatula cDNA clone pgPOD-5D24 5' end,
 mRNA sequence.
 ACCESSION BI308419
 VERSION BI308419.1 GI:14982746
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases: 1 to 804)
 Grusak, M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T., Cho
 J., and Fraser, C.M.
 ESTs from developing reproductive tissues of Medicago truncatula
 Unpublished (2001)
 Contact: Michael A. Grusak
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA
 Tel: 713-798-7044
 Fax: 713-798-7078
 Email: mgrusak@bcm.tmc.edu
 B395274e
 TIGR sequence name: MTOAP24TK
 More information is available at: www.medicago.org
 Seq primer: SKmod (CTA gaa cta gta gAT CC).
 Location/Qualifiers
 1..804
 /organism="Medicago truncatula"
 /cultivar="Al7"
 /db_xref="taxon:3880"
 /clone="pgPOD-5D24"
 /clone_lib="GP0D"
 /tissue_type="immature pod walls"
 /dev_stage="immature pods, ranging in age from 15 to 30
 days after pollination"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Immature pods, ranging in age from 15 to 30 days
 after pollination, were collected from greenhouse-grown
 plants. At harvest, seeds were removed from pods and
 isolated pod walls were collected and immediately frozen
 in liquid nitrogen. Pod walls were pooled for mRNA
 extraction. cDNA was prepared from polyA+ enriched RNA.
 The cDNA was directionally ligated into the Unizap XR
 vector from Strategene and packaged using Gigapack III
 Gold packaging extracts. Plasmids containing cDNA inserts
 were excised from the recombinant lambda-zap phage using
 Ex-assist helper phage and propagated in XL0R cells."
 228 a 167 c 137 g 272 t

Query Match 12.9%; Score 182; DB 13; Length 804;
 Best Local Similarity 70.8%; Pred. No. 3.4e-31;
 Matches 242; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 578 GTCTGTTTCTTCAGATGCAATTCCTACATAAAGGCTATGCGTTTGCAAAATGTTGTG 637
 241 GTTTATGCTCCATAAAGATGCCATTCATGCAATGATGTTCTTTCATAATATCGTTT 300

BASE COUNT 228 a 167 c 137 g 272 t
 ORIGIN

JOURNAL
COMMENT

Unpublished (2000)
Contact: Cathy Ranning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES

source

Location/Qualifiers

1. .722
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone_lib="CSTS20L7"
/clone="CSTS20L7"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

BASE COUNT 191 a 165 c 148 g 218 t

ORIGIN

Query Match 12.1%; Score 170.2; DB 12; Length 722;

Best Local Similarity 69.3%; Pred. No. 1.7e-28;

Matches 232; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 578 GTCGTGTTTCAGATCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTGTG 637

Db 193 GTCTATATCCCAAGATGCATACATCAAGGAAGAGTCTCTTGCAAAATATCAGTT 252

Qy 638 GCAATGAGGCTATGCGATGCTACACTCTCTCCAACTCTCCGAGAGTATGATGAA 697

Db 253 GCTATGAAGCTATGCGGTGTACTGTGCCCTTCCATCACTTTCTGAATACATGATGAA 312

Qy 698 CGTGGTTGGACCAAAATGTTTGTAGCATAGACGAATTCGGCTGGATTCGTATTGTT 757

Db 313 AACGGATGGACCAAAATGTTTGGAGATGAATGATGTTGGATGGCTACTACATCATC 372

Qy 758 TACATGCCATCATCTGTTTTCGTTAGTTGGTATTTATTTGGATGCACAGAGCTT 817

Db 373 AATGCGCTATTTATCTGTAATGAGTGGAGTTTGAATCTACTGATGCACAAAGTTGTTG 432

Qy 818 CATGACATTAAGCTCTCTATAAGTATCTCCATGCCACCATCATATATTTACAACAGCA 877

Db 433 CATGACATAAAACCTCTGTACAATATCTGATGCTACACATCATATTTACAACAGCAA 492

Qy 878 AATACATCTCTCCATTTGCCGGTAAGTGTGTTTCA 912

Db 493 AACACACTTTCCCGTTGCTGGATTGGCATTCCA 527

RESULT 11

BQ803827

LOCUS

WHE2842_D06_G12S Triticum monococcum vernalized apex cDNA library
Triticum monococcum cDNA clone WHE2842_D06_G12, mRNA sequence.

ACCESSION

BQ803827

VERSION

BQ803827.1

KEYWORDS

EST.

SOURCE

Triticum monococcum.

ORGANISM

Triticum monococcum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae

; Triticeae; Triticum.

1 (bases 1 to 574)

Anderson, O.D., Chao, S., Crossman, C., Dubcovsky, J., Echenique, V.,

Lazo, G.R., Pham, J., Rausch, C.J., Stamova, B., Wilson, C. and Woo, J.

The structure and function of the expressed portion of the wheat

genomes - Vernalized apex cDNA library from Triticum monococcum

Unpublished (2002)

COMMENT

Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES

source

Location/Qualifiers

1. .574
/organism="Triticum monococcum"
/cultivar="G3116"
/db_xref="taxon:4568"
/clone="WHE2842_D06_G12"
/clone_lib="Triticum monococcum vernalized apex cDNA
library"
/tissue_type="Vernalized apex"
/dev_stage="One month old plants"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; One-month old plants were
subjected to vernalization treatment by placing them in
the cold room at 6 C, under 15hr light/9hr dark condition.
Total RNA was prepared from apex tissue extracted from
plants with no cold treatment; and from plants with 2-week
, 4-week and 6-week cold treatment separately. Equal
amount of total RNA was pooled from all four samples, a
cDNA library was made using pooled polyA RNA and cDNA
clones were in vivo excised at the University of
California, Davis (V. Echenique, B. Stamova, J. Dubcovsky
) . Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."

BASE COUNT

ORIGIN

Query Match 12.0%; Score 169.6; DB 14; Length 574;
Best Local Similarity 70.6%; Pred. No. 2.2e-28;
Matches 226; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 592 AGATGCAATTCCTACAATAAGGCTATGCGTTTGCAAAATGTTTGGCAATGAAGGCTAT 651

Db 231 AGATGTCGTCCTACAGTAGAAGCTATGAAGCAAAATATTTGTCATCAAAAGGCTAT 290

Qy 652 GCATGTACACTCTCTTCCAACTGCTCCGAGAGTATGATTCAGCTGTTGGACCAA 711

Db 291 GCCTTTCTACTGTGCTCTTCCATCCGATCTGACACATGATGAGAGTGGATGGACAG 350

Qy 712 ATGTTTGTAGCATAGACGAATTCGGCTGGATTCTGTATTGTTTGTATCGGCATCTA 771

Db 351 GTGTTTCTTCATATACAGCAAGTTGGTTGCCCATGTACATTTGCTATGTCTCTATA 410

Qy 772 TCTGTTTTCGTTGAGTTGTTGTTATTTGATGCACAGAGAGCTTCATGACATTAAGCC 831

Db 411 TCTCATCTTTGAGTTTGGATTCTGATGCACAGAGAGTTGTCATACATAAAGCC 470

Qy 832 TCTCTAAGTATCTCCATGCCACCATCATATCTACAACAGCAGATAACACTCTCTCC 891

Db 471 ACTATACAGCATCTACATGCAACCACCATTTTACAACAGGAGATAACCTATCACC 530

Qy 892 ATTTGCGGTAAGTGTGTTTC 911

Db 531 ATTTGCGTGGACTAGCATTC 550

RESULT 12

BE602261

LOCUS

BE602261 708 bp mRNA linear EST 22-OCT-2001
HVSMEH0098E01f Hordeum vulgare 5-45 DAP spike EST library

DEFINITION HVCDA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0098E01f,
mRNA sequence.

ACCESSION BE602261

VERSION BE602261.2 GI:13189996
 KEYWORDS EST.
 SOURCE Hordeum vulgare.
 ORGANISM Hordeum vulgare.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.

AUTHORS Wing, R., Close, T.J., Kleinjohs, A., Wise, R., Begum, D., Frisch, D., Yu
 , Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Penton
 , R.D., Close, S.J., Oates, R. and Main, D.

TITLE Development of a genetically and physically anchored EST resource
 for barley genomics: Morex 5-45 DAP spike cDNA library

JOURNAL Unpublished (2001)

COMMENT On Aug 21, 2000 this sequence version replaced gi:9859822.

Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 375

Seq primer: ATTTACCTCTACTAAAGG

High quality sequence start: 7

High quality sequence stop: 698.

Location/Qualifiers

1. .708

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone="HVSMEH009BE01F"

/clone_lib="Hordeum vulgare 5-45 DAP spike EST library

HVCNDA0009 (5 to 45 DAP)"

/tissue_type="5-45 DAP Spike"

/lab_host="SOLR"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
 Plants were grown in the greenhouse at the University of
 California, Riverside (Fenton, SJ Close, TJ Close). Whole
 spikes with awns trimmed were collected at 5, 10, 15, 20,
 30 and 45 DAP (Fenton). Total RNA was prepared from each
 pool, equal quantities of all six RNA pools were combined,
 poly(A) RNA was purified from the mixture, one primary
 unamplified cDNA library was made, and 1 million pfu were
 in vivo excised to give pBluescript SK(-) cDNA phagemids
 (Choi) in the TJ Close lab at the University of California,
 Riverside. Phagemids were plated and picked at the Clemson
 University Genomics Institute (CUGI) (Begum, Palmer,
 Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
 sequencing and sequence analysis were performed at CUGI
 (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
 The sequence has been trimmed to remove vector sequence
 and contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see close TJ, Wing R, Kleinjohs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/Dgn/31/cover.html)"

BASE COUNT 144 a 210 c 182 g 169 t

Query Match 12.0%; Score 168.8; DB 10; Length 708;

Best Local Similarity 71.5%; Pred. No. 3.5e-28;

Matches 221; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 592 AGATGCAATTCCTACAATAAGGCTATGCTTTCGCAATTTGTTGGCAATCAAGGCTAT 651

Db 301 AGATGCTGCTCCTACAGTAGAGCTATGAGAAGCAATATTTGTCATCAAGGCTAT 360

QY 652 GCCATGGTACACTCTTCTCCAACTGCTCCGAGAGATGATGTAACGTGGTGGACCA 711

Db 361 GCCTTTCTACTGCGCTCTTCCTGCTGATCTGAGCACATGATGAGAGTGGATGGACCG 420
 QY 712 ATGTTTTCAGCATGACGAGAAATTCGGCTGGATTCTGTATTTTGTATACATGCCCTA 771
 Db 421 GTGTTTTTTCATATCAGTAGAAGTTGGTTGGCCGATGATTTTGTATATGTGGTWTATA 480
 QY .772 TCTGTTTTTCGTTGGATTGTTGTTATTTATTTGGATGCACAGAGAGCTTCATGACATTAAGCC 831
 Db 481 TCTACCTTTGTTGGAGTTTGAATTTACTGGATGCACAGAGAGTTGATGACATAAACC 540
 QY 832 TCTCTATAGTATCTCCATGCCACCATCATATATACAAAGCAGCATACACTCTCTCC 891
 Db 541 ACTGTATAGCACCTACACGCCAACCCACCATTTTACAAAGGAGATACCTATCACC 600
 QY 892 ATTTGCGG 900
 Db 601 ATTTGCTGG 609

RESULT 13

BQ472907

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BQ472907 612 bp mRNA linear EST 31-MAY-2002
 HB01C02r BC Hordeum vulgare cDNA clone HB01C02 5-PRIME, mRNA
 sequence.

BQ472907

BQ472907

EST.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

Hordeum vulgare.

REFERENCE 1 (bases 1 to 612)
 AUTHORS Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.
 TITLE Barley ESTs from developing seeds
 JOURNAL Unpublished (2002)
 COMMENT Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 612 Std Error: 0.00
 Plate: 1 row: C column: 2
 Seq primer: M13rev.

FEATURES source

1. .612

/organism="Hordeum vulgare"

/cultivar="Barke"

/db_xref="taxon:4513"

/clone="HB01C02"

/clone_lib="BC"

/tissue_type="developing caryopsis"

/dev_stage="8-15 DAP (days after pollination)"

/lab_host="Xl10-Gold"

/note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of

cDNA); Site_2: XhoI (3'-end of cDNA); developing caryopsis

, 8-15 DAP(days after pollination) due to a cloning

artefact caused by the kit, in most cases the EcoRI site

is NOT present, as well as the EcoRI adapter used for

cloning. To excise the insert, restriction sites upstream

EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also

due to the cloning system used Blue/white selection for

recombinants is not 100% reliable."

BASE COUNT 133 a 169 c 143 g 167 t

ORIGIN

Query Match

Best Local Similarity

Matches 221; Conservative

0; Mismatches 88; Indels

0; Gaps

0;

592	AGATG:AAATCTACAAATAAGGCTATCCGTTTGCAAAATGTTGTGGCAATGAAGGCTAT	651
199	AGATG:ATGCCCCACAGTAGAAGCTATGAAGAAGCAAAATAATTGTCATCAAAAGTCTAT	258
652	GCATAGTACACTCTTTCCAACTGTCTCCGAGAGTATGATTGAACGTGTTGGACCAA	711
259	GCGTTCTACTGTGCTCTCTCTCTGTATCTGAGCACATCATTCAGAGTGGATGGACACG	318
712	ATGTTTTCGTAGCATAGACGAATTCGCTGGATTCTGTAATTTGTTTACATCGCCATCTA	771
319	GTGTTTTTTTCATATCAGTGAAGTTGGTTGGCGATGTATTTGTATATGTGCTTTATA	378
772	TCTTGTTTTCGTGTAGTTTGGTATTTATTGGATGCCACAGAGAGCTTCATGACATTAAAGCC	831
379	TCTCACTTTTGTGTGGAGTTTGTACTGGATGCCACAGAGAGTTGCATGACATAAAACC	438
832	TCTCTATTAAGTATCTCCATGCCACCCATCATATCTACAACAAGCAGAAATACACTCTCTCC	891
439	ACTGTATTAAGCACTACATGCACCCACCACATTTTACAACAAGGAATACCCCTATCACC	498
892	ATTTG:CGG	900
499	ATTTG:TTGG	507

RESULT 14					
BQ802075					
LOCUS					
DEFINITION	BQ802075;	769 bp	mRNA	linear	Est 30-JUL-2002
	WHE2821.G07_M13Z5	Triticum monococcum	vernalized apex	cdna library	
	Triticum monococcum	cdna clone	WHE2821.G07_M13	mRNA sequence.	

cDNA library was made using pooled polyA RNA and cDNA clones were in vivo excised at the University of California, Davis (V. Echenique, B. Stamova, J. Dubcovsky). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

[illegible]

RESULT 15					
BG098248					
LOCUS	BG098248	597 bp	mrna	linear	EST 29-JAN-2001
DEFINITION	EST462767	sprouting eyes/shoots Solanum tuberosum cDNA clone			
	CSTC2K9	5' sequence, mRNA sequence.			

FEATURES	TITLE	, B.
source	Location/Qualifiers	Generation of ESTs from potato sprouting eyes/shoots
	1..769	Unpublished (2001)
	/organism="Triticum monococcum"	Contact: Cathy Ronning
	/cultivar="G3116"	The Institute for Genomic Research
	/db_xref="taxon:4568"	For clone info: please contact Research Genetics, Libraries
	/clone="WHE2821_G07_M13"	Division tel 1-800-711-6195, email cdna@resgen.com.
	/clone_lib="Triticum monococcum vernalized apex cDNA library"	Location/Qualifiers
		1..597
	FEATURES	source

FEATURES source

the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers.^a

BASE COUNT	155 a	138 c	125 g	179 t	
ORIGIN					
Query Match	11.9%	Score 168:	DB 12:	Length 597:	
Best Local Similarity	70.3%	Pred. No. 5.1e-28:			
Matches 225:	Conservative 0:	Mismatches 95:	Indels 0:	Gaps 0:	
Qy	578	GTCTGTTTTCTTCAGATGCAATTCCTACAAATAAAGGCTATGGTTTGTGCAAAATGTTTGTG	637		
Db	278	GTCTATATCCCAAGATGCCATACCATCAAGGAAGCAATGCTCTTGCAAATATCAAGTT	337		
Qy	638	GCAATGAAGGCTATGCCATGGTACACACTCTTTCCAACTCTCTCCGAGAGATGATGAA	697		
Db	338	GCTATGAAGACTATGCGGTGGTACTGTGCGCTTCCATCACTTCTGTAATACATGATGAA	397		
Qy	698	CGTGGTTGGACCAAAATGTTTTGCTAGCATAGACGAATTCGGCTGGGATTCGTATTTGTT	757		
Db	398	AACGGATGGACCAAAATGTTTTGGGAGAATTAAGTGATGTTGGATGGCCCTACCTACATCATC	457		
Qy	758	TACATGCCCATCTATCTGTTTTCGTTGAGTTGGTATTTATTGGATGCACAGAGCITT	817		
Db	458	AATGGCGCTATTATCTTGTAATAGTGAGATTGGGAATCTACTGGATGCACAAGTTGTTG	517		
Qy	818	CATGAGATTAAAGCCTCTCTATAAGTATCTCCATGCCACCCCATCATATCTTACAACAAGCAG	877		
Db	518	CATGACATAAAACCTCTGTACAANAATCTGCAATGCTACACATCATATTTACACACAGCAA	577		
Qy	878	AATACACTCTCTCCATTTGC	897		
Db	578	AACACACTTTCCCGGTTTC	597		

Search completed: June 6, 2003, 16:23:13
Job time : 3670.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 10:02:20 ; Search time 54 Seconds
(without alignments)

7506.661 Million cell updates/sec

Title: US-09-775-879-20_COPY_143_322

Perfect score: 180

Sequence: 1 atgcggcggaatgtta.....tcattcgggttctctg 180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185139 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	100.0	1164	21 AAC41897	Arabidopsis thalia
2	54.8	30.4	1381	22 AAH50964	Sterol C5 desaturase
3	32.6	18.1	1830121	17 AAT42063	Haemophilus influenzae
C 4	31.8	17.7	1283	22 AAK66079	Human immune/haema
C 5	31.8	17.7	1283	22 AAK66083	Human immune/haema
C 6	31.8	17.7	1283	22 AAK66084	Human immune/haema
C 7	30.2	16.8	311	22 AAH36614	Human colon cancer
C 8	30.2	16.8	2667	24 AAH69295	Listeria innocua D
C 9	30.2	16.8	2709	24 ABQ67863	Listeria innocua D

C 10	30.2	16.8	81905	24 ABQ69244	Listeria innocua D
C 11	30.2	16.8	82689	24 ABQ67198	Listeria innocua p
C 12	30	16.7	4775	22 AAK76537	Human immune/haema
C 13	30	16.7	4776	22 AAK76538	Human immune/haema
C 14	30	16.7	11316	22 AAS27731	DNA encoding novel
C 15	29.6	16.4	773	22 AAH03702	Human CDNA clone (
C 16	29.6	16.4	919	22 AAL37315	Human musculoskele
C 17	29.6	16.4	2310	22 AAH15301	Human CDNA sequenc
C 18	29.6	16.4	3810	24 ABN97214	Gene #3712 used to
C 19	29.2	16.2	1403	20 AAX86970	Truncated plastid
C 20	29.2	16.2	2381	24 ABN59690	Novel human coding
C 21	29.2	16.2	2703	21 AAN70093	Plasmodium falcipa
C 22	28.8	16.0	2114	14 AAQ39178	Truncated PVX repl
C 23	28.8	16.0	2208	14 AAQ39180	Truncated PVX repl
C 24	28.8	16.0	4455	14 AAQ39177	PVX ORF1 encoding
C 25	28.6	15.9	1884	23 ABL13343	Drosophila melanog
C 26	28.4	15.8	1638	20 AAX16150	NADH-H2O oxidase a
C 27	28.2	15.7	87350	18 AAX83003	Human WRN genomic
C 28	28	15.6	2192	24 AAD22130	Arabidopsis thalia
C 29	28	15.6	3647	23 ABL02676	Drosophila melanog
C 30	28	15.6	35100	20 AAV73804	KSHV LTR DNA (nucl
C 31	28	15.6	37544	21 AAA50029	Cosmid cHRIM5 enco
C 32	28	15.6	137507	19 AAV19941	KSHV long unique c
C 33	28	15.6	1664976	19 AAV21209	Methanococcus jann
C 34	27.8	15.4	1977	22 AAH68280	C glutamicum codin
C 35	27.8	15.4	309400	22 AAH68534	C glutamicum codin
C 36	27.6	15.3	805	21 AAC77362	Human ORFX ORF291?
C 37	27.6	15.3	820	22 AAH05833	Human CDNA clone (
C 38	27.6	15.3	1003	22 AAH08426	Human guanine exch
C 39	27.6	15.3	1003	22 AAK52823	Human polynucleoti
C 40	27.6	15.3	1210	22 AAI60003	Human polynucleoti
C 41	27.6	15.3	1897	22 AAK51839	Human polynucleoti
C 42	27.6	15.3	1897	22 AAI58217	Human polynucleoti
C 43	27.6	15.3	2276	21 AAA59498	DNA encoding Smad1
C 44	27.6	15.3	2397	22 AAH14389	Human CDNA sequenc
C 45	27.6	15.3	2720	22 AAS33176	DNA encoding human

ALIGNMENTS

RESULT 1

ID AAC41897 standard; DNA; 1164 BP.

XX AC AAC41897;

XX AC AAC41897;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33540.

XX DE Arabidopsis thaliana

XX DE Hybridisation assay; genetic mapping; gene expression control;

XX DE Protein identification; signal transduction pathway;

XX DE Metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142320.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144321.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 90S-0159330.
 PR 14-OCT-1999; 90S-0159331.
 PR 14-OCT-1999; 90S-0159637.
 PR 14-OCT-1999; 90S-0159638.
 PR 18-OCT-1999; 90S-0159584.
 PR 21-OCT-1999; 90S-0160741.
 PR 21-OCT-1999; 90S-0160767.
 PR 21-OCT-1999; 90S-0160768.
 PR 21-OCT-1999; 90S-0160770.
 PR 21-OCT-1999; 90S-0160814.
 PR 21-OCT-1999; 90S-0160815.
 PR 22-OCT-1999; 90S-0160980.
 PR 22-OCT-1999; 90S-0160981.
 PR 22-OCT-1999; 90S-0160989.
 PR 25-OCT-1999; 90S-0161404.
 PR 25-OCT-1999; 90S-0161405.
 PR 25-OCT-1999; 90S-0161406.
 PR 26-OCT-1999; 90S-0161359.
 PR 26-OCT-1999; 90S-0161360.
 PR 26-OCT-1999; 90S-0161361.
 PR 28-OCT-1999; 90S-0161920.
 PR 28-OCT-1999; 90S-0161922.
 PR 28-OCT-1999; 90S-0161993.
 PR 29-OCT-1999; 90S-0162142.

 Query Match 100.0%; Score 180; DB 21; Length 1164;
 Best Local Similarity 100.0%; Pred. No. 3.9e-53;
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 ATGGCGCGGATAAGCTTAATCTGATGAGTGTGTGAGGAACCTCTTTTACACCGA 60
 Db 146 ATGGCGCGGATAAGCTTAATCTGATGAGTGTGTGAGGAACCTCTTTTACACCGA 205

 -QY 61 ATCGTTTGAGTCATCTTTGCGCGGGAATCTATGGAACCTTACCTCATTTCTCCAG 120
 Db 206 ATCGTTTGAGTCATCTTTGCGCGGGAATCTATGGAACCTTACCTCATTTCTCCAG 265

 QY 121 ACATGGTCCGAAATACCTCCCGGAACCTTACTATCTTCTATCTTCTCGGTTTCTCTGG 180
 Db 266 ACATGGTCCGAAATACCTCCCGGAACCTTACTATCTTCTATCTTCTCGGTTTCTCTGG 325

 RESULT 2
 AAH50964
 ID AAH50964 standard; cdNA; 1381 BP.
 XX AC AAH50964;
 XX AC
 DT 28-AUG-2001 (first entry)
 XX AC
 DE Sterol C5 desaturase long partial clone nucleotide sequence.
 XX AC
 KW Moss; Physcomitrella patens; lipid metabolism related protein; LMRP;
 KW lipid biosynthesis; lipid modification; lipid degradation; cofactor;
 KW fatty acid transport; genetic engineering; fatty acid; enzyme; plant;
 KW microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;
 KW biotic stress tolerance; abiotic stress tolerance; rye; oat; triticale;
 KW rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot;
 KW pepper; sunflower; tagetes; potato; tobacco; eggplant; tomato; Vicia;
 KW pea; alfalfa; coffee; cacao; tea; Salix; oil palm; coconut;
 KW perennial grass; forage crop; ss.
 XX AC
 OS Physcomitrella patens.
 XX AC
 PN WO200138484-A2
 XX AC
 PD 31-MAY-2001.
 XX AC
 PF 22-NOV-2000; 2000WO-EP11615.
 XX AC
 PR 25-NOV-1999; 99WO-EP09108.
 XX AC
 PA (BADI) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R;
 XX WPI; 2001-367669/38.
 DR Nucleic acids encoding lipid metabolism related proteins from
 XX Physcomitrella patens useful to produce fine chemicals in modified
 PT organisms, particularly polyunsaturated fatty acids in oilseed plants -
 PT Claim 7; Page 106; 120pp; English.
 PS The present invention describes isolated nucleic acid sequences which
 XX encode lipid metabolism related proteins (LMRP). The LMRP nucleic acids
 CC can be used to modify lipids and fatty acids, cofactors and enzymes in
 CC microorganisms and plants, particularly to produce polyunsaturated fatty
 CC acids, and are especially useful in oilseed plants. The nucleic acids
 CC may also confer biotic or abiotic stress tolerance, particularly to
 CC maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut,
 CC cotton, rapeseed, canola, manihot, pepper, sunflower, tagetes, potato,
 CC tobacco, eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao,
 CC tea, Salix species, oil palm, coconut, perennial grasses and forage
 CC crops. AAH50878 to AAH50882 represent primers used in the exemplification
 CC of the present invention. AAH50883 to AAH50968 represents LMRP nucleotide
 CC sequences, and AAH50843 to AAH50928 represent LMRP protein sequences,
 CC given in the present invention.
 XX SQ Sequence 1381 BP; 323 A; 331 C; 351 G; 376 T; 0 other;

 Query Match 30.4%; Score 54.8; DB 22; Length 1381;
 Best Local Similarity 59.7%; Pred. No. 5.1e-09;
 Matches 92; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

 QY 27 GCAGTTTGTTGAGGAACCTCTTTTACACCGAATCTGAGTCATCTTTGCGCGG 86
 Db 231 GCTTTTCGTGGAGGAGACCCCGGTGATCAACGATCTGGTGGCCCTGGCTGCCCTC 290

 QY 87 GAATCTATGGGAACCTTACCTCATTTCTCCAGACATGGCTCCGAAATACCTGCCCGG 146
 Db 291 CTCTGTCGCGGACTCCATTCGCCACACATTCGAGACATGGCTGCGGAACCTAGCTGCGGG 350

 QY 147 AACCTTACTATCTTCACTCCGGTTTCTCTGG 180
 Db 351 CATGCTTTGTATTTCTGCTCCGCGTGGCTGTGG 384

 RESULT 3
 AAT42063
 ID AAT42063 standard; DNA; 1830121 BP.
 XX AC AAT42063;
 XX AC
 DT 14-SEP-1999 (first entry)
 XX AC
 DE Haemophilus influenzae complete genome sequence.
 XX AC
 KW Genome; bacterium; Haemophilus influenzae; computer readable medium;
 KW expression modulating fragment; regulation; gene expression; vector;
 KW organism; open reading frame; ORF; ds.
 XX AC
 OS Haemophilus influenzae.
 XX AC
 PN WO9633276-A1.
 XX AC
 PD 24-OCT-1996.
 XX AC
 PF 22-APR-1996; 96WO-US05320.
 XX AC
 PR 07-JUN-1995; 95US-0487429.
 XX AC
 PR 21-APR-1995; 95US-0426787.
 XX AC
 PR 07-JUN-1995; 95US-0476102.
 XX AC
 PA (HUMA-) HUMAN GENOME SCI INC.

PA (UYJO) UNIV JOHNS HOPKINS.
XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
XX WPI; 1996-485782/48.
XX Haemophilus influenzae Rd genome recorded on computer readable
PT medium - useful for identifying commercially important nucleic acid
PT fragments by homology searching
XX
PS Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
CC This sequence represents the complete genome sequence of the bacterium
CC Haemophilus influenzae strain Rd. The invention relates to a computer
CC readable medium (CRM) having recorded upon it the complete H.influenzae
CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC sequence at least 99% identical to (I). By providing the full-length
CC genomic sequence in a computer readable form, it is possible to identify
CC commercially important nucleic acid fragments and expression modulating
CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC regulate the expression of a nucleic acid molecule. Vectors and altered
CC organisms comprising the predicted ORFs can be used to produce any of the
CC polypeptide fragments of the H. influenzae Rd genome.
XX
SQ Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Query Match 18.1%; Score 32.6; DB 17; Length 1830121;
Best Local Similarity 54.6%; Pred. No. 5.3;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 59 GAATCGTTCGAGTCATCTTTGCGCGGCGGAATCTATGGAACCCCTTACCTCATTTTCCTCC 118
DB 1431803 GAGCCATACAGACATCTTTGGTGCAGAAATACCCGACATCATATTTATTACCCA 1431862

QY 119 AGACATGGCTCCGAATTAACCTCGCGGAACCTACTACTCTTCCTCGGTTTCCTC 177
DB 1431863 AATTATTGAGAGATTAGCTTCGGGTAAACATAAAGCGTTATACCCGTTTCCCG 1431921

RESULT 4
AAK66079/c
ID AAK66079 standard; DNA; 1283 BP.
XX
AC AAK66079;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20891.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0228279.
PR 22-AUG-2000; 2000US-0228681.
PR 22-AUG-2000; 2000US-0228688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240560.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in RAM921170 to RAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting

Db 783 GGATTAAGATTATTAATGCTGCTGTAAGGGAACCACTTTTCATACACAAATGTATC 724
QY 69 GAGTCATCTTTGCGGCGAATCTATGGGAACCCCTTACCTCATTTTCTCCAGACA 123
Db 723 TTTTTCCTTTTCARAGGCTATACACAGTACATTTACATTAACCTAATCCCAACA 669

RESULT 6
AAK66084/c
ID AAK66084 standa:d; DNA; 1283 BP.
XX AC AAK66084;
XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20896.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226688.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 21-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 25-SEP-2000; 2000US-0234274.
XX PR 25-SEP-2000; 2000US-0234997.
XX PR 25-SEP-2000; 2000US-0234998.
XX PR 26-SEP-2000; 2000US-0235484.
XX PR 27-SEP-2000; 2000US-0235834.
XX PR 27-SEP-2000; 2000US-0235836.
XX PR 29-SEP-2000; 2000US-0236327.
XX PR 29-SEP-2000; 2000US-0236367.
XX PR 29-SEP-2000; 2000US-0236368.
XX PR 29-SEP-2000; 2000US-0236369.
XX PR 29-SEP-2000; 2000US-0236370.
XX PR 02-OCT-2000; 2000US-0236802.
XX PR 02-OCT-2000; 2000US-0237037.
XX PR 02-OCT-2000; 2000US-0237038.
XX PR 02-OCT-2000; 2000US-0237039.
XX PR 13-OCT-2000; 2000US-0237040.
XX PR 13-OCT-2000; 2000US-0239935.
XX PR 20-OCT-2000; 2000US-0239937.
XX PR 20-OCT-2000; 2000US-0240960.
XX PR 20-OCT-2000; 2000US-0241221.
XX PR 20-OCT-2000; 2000US-0241785.
XX PR 20-OCT-2000; 2000US-0241786.
XX PR 20-OCT-2000; 2000US-0241808.
XX PR 20-OCT-2000; 2000US-0241809.
XX PR 20-OCT-2000; 2000US-0241826.
XX PR 01-NOV-2000; 2000US-0244617.
XX PR 08-NOV-2000; 2000US-0246474.
XX PR 08-NOV-2000; 2000US-0246475.
XX PR 08-NOV-2000; 2000US-0246476.
XX PR 08-NOV-2000; 2000US-0246477.
XX PR 08-NOV-2000; 2000US-0246478.
XX PR 08-NOV-2000; 2000US-0246523.
XX PR 08-NOV-2000; 2000US-0246524.
XX PR 08-NOV-2000; 2000US-0246525.
XX PR 08-NOV-2000; 2000US-0246526.
XX PR 08-NOV-2000; 2000US-0246527.
XX PR 08-NOV-2000; 2000US-0246528.
XX PR 08-NOV-2000; 2000US-0246532.
XX PR 08-NOV-2000; 2000US-0246532.
XX PR 08-NOV-2000; 2000US-0246610.
XX PR 08-NOV-2000; 2000US-0246611.
XX PR 08-NOV-2000; 2000US-0246613.
XX PR 17-NOV-2000; 2000US-0249207.
XX PR 17-NOV-2000; 2000US-0249208.
XX PR 17-NOV-2000; 2000US-0249209.
XX PR 17-NOV-2000; 2000US-0249210.
XX PR 17-NOV-2000; 2000US-0249211.
XX PR 17-NOV-2000; 2000US-0249212.
XX PR 17-NOV-2000; 2000US-0249213.
XX PR 17-NOV-2000; 2000US-0249214.
XX PR 17-NOV-2000; 2000US-0249215.
XX PR 17-NOV-2000; 2000US-0249216.
XX PR 17-NOV-2000; 2000US-0249217.
XX PR 17-NOV-2000; 2000US-0249218.
XX PR 17-NOV-2000; 2000US-0249244.
XX PR 17-NOV-2000; 2000US-0249245.


```
PR 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
PI
XX WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX
XX Claim 5; SEQ ID 11; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 81905 BP; 26162 A; 14249 C; 14844 G; 26650 T; 0 other;
SQ
Query Match 16.8%; Score 30.2; DB 24; Length 81905;
Best Local Similarity 51.9%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 11 ATAATGCTTATCTGATGCGAGTTTGTGACGAAACCTCTTTTACAAACCGAATCGTTCTGA 70
DB 49965 ATATTGATATTAATCTGTCATAATTAATTGAGAAATCAATATTACATAAGATCCCTCATCT 49906
QY 71 GTCATCTTTTGGCGGGAATCTATGGAACCCCTTACCTCATTTTCTCCAGACATGGCTCC 130
DB 49905 TGAATCTCCATCAATAATATTATTCGATGCTTTTCTCTATATATGCTTAAGTGTCTCC 49846.
QY 131 GAAATTAACCTC 141
DB 49845 GAAATTAACCTC 49835
RESULT 11
AB067198/c
XX AB067198 standard; DNA; 82689 BP.
XX
XX AC AB067198;
XX
XX 29-AUG-2002 (first entry)
XX
XX Listeria innocua plasmid DNA sequence.
DE
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria innocua.
OS
XX WO200228891-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-FR03061.
XX
XX 04-OCT-2000; 2000FR-0012697.
XX
XX (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
PI
XX WPI; 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
XX treatment and prevention of infection, also related polypeptides,
XX antibodies and modulators
XX
XX Claim 5; SEQ ID 2057; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
XX (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
XX and primers for identification and/or detection of Listeria (e.g. as
XX contaminants in foods, or mutational analysis) and for analysis of
XX gene expression. Proteins encoded by the nucleic acid sequences can be
XX used to screen for compounds that modulate gene expression, replication
XX and pathogenicity of Listeria (potential therapeutic agents), also for
XX treating infections by Listeria, and are useful as immunogens in
XX anti-Listeria vaccines.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 82689 BP; 26448 A; 14348 C; 14977 G; 26916 T; 0 other;
SQ
Query Match 16.8%; Score 30.2; DB 24; Length 82689;
Best Local Similarity 51.9%; Pred. No. 11;
Matches 68; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 11 ATAATGCTTATCTGATGCGAGTTTGTGACGAAACCTCTTTTACAAACCGAATCGTTCTGA 70
DB 36680 ATATTGATATTAATCTGTCATAATTAATTGAGAAATCAATATTACATAAGATCCCTCATCT 36621
QY 71 GTCATCTTTTGGCGGGAATCTATGGAACCCCTTACCTCATTTTCTCCAGACATGGCTCC 130
DB 36620 TGAATCTCCATCAATAATATTATTCGATGCTTTTCTCTATATATGCTTAAGTGTCTCC 36561
QY 131 GAAATTAACCTC 141
DB 36560 GAAATTAACCTC 36550
RESULT 12
AAK76537
ID AAK76537 standard; DNA; 4775 BP.
XX
XX AC AAK76537;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:31349.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX 04-FEB-2000; 2000US-0180628.
XX
XX 24-FEB-2000; 2000US-0184564.
XX
XX 16-MAR-2000; 2000US-0186350.
XX
XX 02-MAR-2000; 2000US-0189874.
XX
XX 17-MAR-2000; 2000US-0190076.
XX
XX 18-APR-2000; 2000US-0198123.
XX
XX 19-MAY-2000; 2000US-0205515.
XX
XX 07-JUN-2000; 2000US-0209467.
XX
XX 28-JUN-2000; 2000US-0214886.
```

PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246539.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC	
XX		
PI	Rosen CA, Barash SC, Ruben	
XX		
DR	WPT; 2001-483426/52.	
XX		
PT	Nucleic acids encoding human	
PT	useful for preventing, diagn	
PT	metastasis -	
XX		
PS	Disclosure; SEQ ID NO 31349;	
XX		
CC	AAK54951 to AAK64702 encode	
CC	amino acid sequences given i	
CC	activity, and can be used in	
CC	proteins and polynucleotides	
CC	treatment of diseases associ	
CC	example, they may be used to	
CC	expression by rectifying mut	

Nucleic acids encoding human immune/hematopoietic antigen polypeptides; useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 31349; 3071pp + Sequence Listing; English.

AAK34951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytotoxic activity, and can be used in gene therapy and vaccine production. (I) proteins, and polynucleotides may be used in the prevention, diagnosis a treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 4775 BP; 1140 A; 1070 C; 974 G; 1591 T; 0 other;

Query Match 16.7%; Score 30; DB 22; Length 4775;
Best Local Similarity 59.3%; Pred. No. 4.4;
Matches 51; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 8 CGGATATCGTTATCGTATGCGAGTTGTTGACGAAACCTCTTTTACACCGCAATGTTTC 67
DB 1283 CTGATATCGTTATGTTATGTCATATGTTGGAAGAAATGTCCTTTCTGTCTAATATAGC 1342
QY 68 TGAGTCATCTTTTCCGGCGGAATCTA 93
DB 1343 TTATATTGTTTATTGTTAATATA 1368

RESULT 13

AAK76538
ID AAK76538 standard; DNA; 4776 BP.

XX AAK76538;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31350.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.

PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 03-SEP-2000; 2000US-0229509.
 PR 03-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231142.
 PR 08-SEP-2000; 2000US-0231143.
 PR 08-SEP-2000; 2000US-0231144.
 PR 08-SEP-2000; 2000US-0231143.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465460/50.

Novel polypeptides useful for diagnosing, treating, preventing and/or
 prognosing disorders related to the proteins, including cancers, immune
 disorders and neuronal disorders -

Claim 1; SEQ ID NO 1391; 880pp; English.

The invention relates to novel isolated polypeptides (I), and
 polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 diagnosing, preventing and treating diseases including immune system
 disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 transplant rejections and graft versus host disease, infectious diseases
 (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 other blood-related disorders (sickle cell anaemia), myeloproliferative
 disorders, primary haematopoietic disorders, hyperproliferative
 disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 disorders (e.g. glomerulonephritis), cardiovascular disorders
 (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 Addison's disease), reproductive system disorders, gastrointestinal
 disorder (inflammatory disorders), liver disorders (cirrhosis),
 as stimulators of B-cell responsiveness to pathogens, activators of
 T-cells, to induce higher affinity antibodies, and as a means to induce
 tumour proliferation in pathologies e.g. acquired immune deficiency
 syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 pathway protein coding sequences and PCR primers of the invention.

Sequence 11316 BP; 2843 A; 2769 C; 2919 G; 2785 T; 0 other;

Query Match 16.7%; Score 30; DB 22; Length 11316;
 Best Local Similarity 61.5%; Pred. No. 6.2;
 Matches .48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

42 AACCTCTTTTACACCGAATCGTCTGAGTCATCTTTCCCGCGCAATATGGGACC 101
 |||| |||| |||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | |||

THIS PAGE BLANK (USPTO)